

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/12, C07K 14/705, 16/28, G01N 33/68, C12N 1/21, 5/10, 1/19, C12Q 1/68, A61K 39/395, 38/17, 31/70		A2	(11) International Publication Number: WO 99/47673 (43) International Publication Date: 23 September 1999 (23.09.99)
(21) International Application Number: PCT/US99/03740 (22) International Filing Date: 16 March 1999 (16.03.99) (30) Priority Data: 09/040,111 17 March 1998 (17.03.98) US (71) Applicant: SCHERING CORPORATION [US/US]; 2000 Galloping Hill Road, Kenilworth, NJ 07033 (US). (72) Inventors: CHALUS, Lionel; 93, route de Paris, F-69570 Dardilly (FR). QUAN, Ahn, B.; 2124 Park Avenue #38, San Jose, CA 95126 (US). BATES, Elizabeth, Esther, Mary; 13, rue Terme, F-69001 Lyon (FR). GORMAN, Daniel, M.; 6371 Central Avenue, Newark, CA 94560 (US). SAELAND, Sem; 6, rue Soeurs Bouvier, F-69005 Lyon (FR). LEBECQUE, Serge, J., E.; 514, chemin du Marand, F-69380 Civrieux d'Azergue (FR). PHILIPPS, Joseph, H., Jr.; 1511 Walnut Drive, Palo Alto, CA 94303 (US). (74) Agent: FOULKE, Cynthia, L.; Schering-Plough Corporation, Patent Dept., K-6-1 1990, 2000 Galloping Hill Road, Kenilworth, NJ 07033-0530 (US).		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GD, GE, HR, HU, ID, IL, IN, IS, JP, KG, KR, KZ, LC, LK, LR, LT, LU, LV, MD, MG, MK, MN, MX, NO, NZ, PL, PT, RO, RU, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UZ, VN, YU, ZA, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>Without international search report and to be republished upon receipt of that report.</i>	
(54) Title: ISOLATED MAMMALIAN MEMBRANE PROTEIN GENES AND RELATED REAGENTS			
(57) Abstract Nucleic acids encoding various dendritic cell membrane proteins, reagents related thereto, including specific antibodies, and purified proteins are described. Methods of using said reagents and related diagnostic kits are also provided.			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	R	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

ISOLATED MAMMALIAN MEMBRANE PROTEIN GENES
AND RELATED REAGENTS

5

FIELD OF THE INVENTION

The present invention contemplates compositions related to genes found in lymphocytes, e.g., cells which function in the immune system. These genes function in controlling development, differentiation, and/or physiology of the mammalian immune system. In particular, the application provides nucleic acids, proteins, antibodies, and methods of using them.

10

15

BACKGROUND OF THE INVENTION

The circulating component of the mammalian circulatory system comprises various cell types, including red and white blood cells of the erythroid and myeloid cell lineages. See, e.g., Rapaport (1987) Introduction to Hematology (2d ed.) Lippincott, Philadelphia, PA; Jandl (1987) Blood: Textbook of Hematology, Little, Brown and Co., Boston, MA.; and Paul (ed. 1993) Fundamental Immunology (3d ed.) Raven Press, N.Y.

20

25

30

35

Dendritic cells (DC) are antigen-processing or presenting cells, and are found in all tissues of the body. They can be classified into various categories, including: interstitial dendritic cells of the heart, kidney, gut, and lung; Langerhans cells in the skin and mucous membranes; interdigitating dendritic cells in the thymic medula and secondary lymphoid tissue; and blood and lymph dendritic cells. Although dendritic cells in each of these compartments are CD45+ leukocytes that apparently arise from bone marrow, they may exhibit differences that relate to maturation state and microenvironment.

These dendritic cells efficiently process and present antigens to, e.g., T cells. They stimulate responses from naive and memory T cells in the

paracortical area of secondary lymphoid organs. There is some evidence for a role in induction of tolerance.

The primary and secondary B-cell follicles contain follicular dendritic cells that trap and retain intact antigen as immune complexes for long periods of time. These dendritic cells present native antigen to B cells and are likely to be involved in the affinity maturation of antibodies, the generation of immune memory, and the maintenance of humoral immune responses.

Monocytes are phagocytic cells that belong to the mononuclear phagocyte system and reside in the circulation. See Roitt (ed) Encyclopedia of Immunology Academic Press, San Diego. These cells originate in the bone marrow and remain only a short time in the marrow compartment once they differentiate. They then enter the circulation and can remain there for a relatively long period of time, e.g., a few days. The monocytes can enter the tissues and body cavities by the process designated diapedesis, where they differentiate into macrophages and possibly into dendritic cells. In an inflammatory response, the number of monocytes in the circulation may double, and many of the increased number of monocytes diapedese to the site of inflammation.

Antigen presentation refers to the cellular events in which a proteinaceous antigen is taken up, processed by antigen presenting cells (APC), and then recognized to initiate an immune response. The most active antigen presenting cells have been characterized as the macrophages, which are direct developmental products from monocytes; dendritic cells; and certain B cells.

Macrophages are found in most tissues and are highly active in internalization of a wide variety of protein antigens and microorganisms. They have a highly developed endocytic activity, and secrete many products important in the initiation of an immune response. For this reason, it is believed that many genes expressed by monocytes or induced by monocyte activation are likely to be important in antigen uptake, processing, presentation, or regulation of the resulting immune response.

However, dendritic cells and monocytes are poorly characterized, both in terms of proteins they express, and many of their functions and mechanisms of action, including their activated states. In particular, the processes and mechanisms related to the initiation of an immune response, including antigen processing and presentation, remain unclear. The absence of knowledge about the structural, biological, and physiological properties of these cells limits their understanding. Thus, medical conditions where regulation, development, or physiology of antigen presenting cells is unusual remain unmanageable.

SUMMARY OF THE INVENTION

The present invention is based, in part, upon the discovery of various mammalian Schering Dendritic Cell Membrane Protein (SDCMP) genes. Distribution data indicates a broader cellular distribution, and structural data suggests some function, and are exemplified by the specific SDCMP3 and SDCMP4 embodiments. The SDCMPs 3 and 4 exhibit similarity to a class of lectins and asialoglycoprotein receptors (ASGPR). The invention embraces agonists and antagonists of the gene products, e.g., mutations (muteins) of the natural sequences, fusion proteins, chemical mimetics, antibodies, and other structural or functional analogs. It is also directed to isolated genes encoding proteins of the invention. Various uses of these different protein or nucleic acid composition are also provided.

In particular embodiments, the invention provides a binding compound comprising an antibody binding site which specifically binds to: a SDCMP3 or SDCMP4 protein. In preferred embodiments, in the binding compound, the antibody binding site is: specifically immunoreactive with a protein of SEQ ID NO: 2, 4, 6, or 8; raised against a purified or recombinantly produced human or rodent SDCMP3 protein; raised against a purified or recombinantly produced human SDCMP4 protein; in a monoclonal antibody, Fab, or F(ab)2; or the binding compound is: detectably labeled; sterile; or in a buffered composition.

The invention embraces methods using those binding compounds, comprising contacting the binding compound with a biological sample comprising an antigen to form a binding compound:antigen complex. In certain
5 embodiments, the biological sample is human or rodent, and the binding compound is an antibody. The invention also provides a detection kit comprising such binding compound and: instructional material for the use of such binding compound for the detection; or a compartment
10 providing segregation of the binding compound.

The invention also provides a substantially pure or isolated polypeptide, which specifically binds to such binding compounds. In various embodiments, the polypeptide: comprises at least a fragment of at least 14
15 amino acid residues from a primate or rodent SDCMP3 protein; comprises at least 14 amino acids from a primate SDCMP4; is a soluble polypeptide; is detectably labeled; is in a sterile composition; is in a buffered composition; binds to an sialic acid residue; is
20 recombinantly produced, or has a naturally occurring polypeptide sequence.

Nucleic acid embodiments are provided, including a nucleic acid encoding a polypeptide above, when purified. Often, the nucleic acid: comprises at least 30
25 nucleotides of the coding portion of SEQ ID NO: 1 or 3; comprises at least 30 nucleotides from the coding portion of SEQ ID NO: 5 or 7; or it may comprise an insert which selectively hybridizes to a nucleic acid encoding a polypeptide defined above. The invention also provides a
30 cell transfected with such a nucleic acid, e.g., which consists of the protein encoding portions of SEQ ID NO: 1, 3, 5, or 7.

The invention provides methods using at least one strand of those nucleic acids to form a duplex nucleic
35 acid, comprising a step of contacting such strand to a sample to a complementary strand capable of specifically hybridizing. In preferred embodiments, the method allows detection of the duplex; or allows histological localization of the duplex.

Alternatively, the invention provides methods of using a described binding composition, comprising a step of contacting the binding composition with a sample to form a binding composition:antigen complex. In preferred
5 embodiments, the sample is a biological sample, including a body fluid; the antigen is on a cell; or the antigen is further purified.

The invention further embraces methods using those polypeptides, comprising contacting the polypeptide with
10 a sample to form a binding composition:polypeptide complex. In preferred embodiments, the polypeptide is further purified.

Another method provided is to modulating dendritic cell physiology or function comprising a step of
15 contacting the cell with: a binding composition as described; a SDCMP3 or SDCMP4 protein as described; or a polypeptide as described. The function may also result in initiation or progression of an immune response. Typically, the contacting is in combination with an
20 antigen, including a cell surface, MHC Class I, or MHC Class II antigen.

DETAILED DESCRIPTION OF THE INVENTION

All references cited herein are incorporated herein
25 by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference in its entirety for all purposes.

30 I. General

The present invention provides DNA sequences encoding mammalian proteins expressed on dendritic cells (DC). For a review of dendritic cells, see Steinman
35 (1991) Annual Review of Immunology 9:271-296; and Banchemreau and Schmitt (eds. 1994) Dendritic Cells in Fundamental and Clinical Immunology Plenum Press, NY. These proteins are designated dendritic cell proteins because they are found on these cells and appear to exhibit some specificity in their expression.

Specific human embodiments of these proteins are provided below. The descriptions below are directed, for exemplary purposes, to human DC genes, but are likewise applicable to structurally, e.g., sequence, related
5 embodiments from other sources or mammalian species, including polymorphic or individual variants. These will include, e.g., proteins which exhibit a relatively few changes in sequence, e.g., less than about 5%, and number, e.g., less than 20 residue substitutions,
10 typically less than 15, preferably less than 10, and more preferably less than 5 substitutions, including 4, 3, 2, or 1. These will also include versions which are truncated from full length, as described, and fusion proteins containing substantial segments of these
15 sequences.

II. Definitions

The term "binding composition" refers to molecules that bind with specificity to a these DC proteins, e.g.,
20 in an antibody-antigen interaction. Other compounds, e.g., proteins, can also specifically associate with the respective protein. Typically, the specific association will be in a natural physiologically relevant protein-protein interaction, either covalent or non-covalent, and
25 may include members of a multiprotein complex, including carrier compounds or dimerization partners. The molecule may be a polymer, or chemical reagent. A functional analog may be a protein with structural modifications, or may be a wholly unrelated molecule, e.g., which has a
30 molecular shape which interacts with the appropriate interacting determinants. The variants may serve as agonists or antagonists of the protein, see, e.g., Goodman, et al. (eds.) (1990) Goodman and Gilman's: The Pharmacological Bases of Therapeutics (8th ed.) Pergamon
35 Press, Tarrytown, N.Y.

The term "binding agent:DC protein complex", as used herein, refers to a complex of a binding agent and DC protein. Specific binding of the binding agent means that the binding agent has a specific binding site that
40 recognizes a site on the respective DC protein. For

example, antibodies raised to the DC protein and recognizing an epitope on the DC protein are capable of forming an antibody:DC protein complex by specific binding. Typically, the formation of a binding agent:DC protein complex allows the measurement of that DC protein in a mixture of other proteins and biologics. The term "antibody:DC protein complex" refers to a binding agent:DC protein complex in which the binding agent is an antibody. The antibody may be monoclonal, polyclonal or even an antigen binding fragment of an antibody, e.g., including Fv, Fab, or Fab2 fragments.

"Homologous" nucleic acid sequences, when compared, exhibit significant similarity. The standards for homology in nucleic acids are either measures for homology generally used in the art by sequence comparison and/or phylogenetic relationship, or based upon hybridization conditions. Hybridization conditions are described in greater detail below.

An "isolated" nucleic acid is a nucleic acid, e.g., an RNA, DNA, or a mixed polymer, which is substantially separated from other components which naturally accompany a native sequence, e.g., proteins and flanking genomic sequences from the originating species. The term embraces a nucleic acid sequence which has been removed from its naturally occurring environment, and includes recombinant or cloned DNA isolates and chemically synthesized analogs or analogs biologically synthesized by heterologous systems. A substantially pure molecule includes isolated forms of the molecule. An isolated nucleic acid will generally be a homogeneous composition of molecules, but will, in some embodiments, contain minor heterogeneity. This heterogeneity is typically found at the polymer ends or portions not critical to a desired biological function or activity.

As used herein, the term "SDCMP3 protein" shall encompass, when used in a protein context, a protein having amino acid sequences as shown in SEQ ID NO: 2 or 4, or a significant fragment of such a protein. It refers to a polypeptide which interacts with the respective SDCMP3 protein specific binding components.

These binding components, e.g., antibodies, typically bind to the SDCMP3 protein with high affinity, e.g., at least about 100 nM, usually better than about 30 nM, preferably better than about 10 nM, and more preferably at better than about 3 nM. Similarly, the use of the term SDCMP4 will apply with reference to SEQ ID NO: 6 or 8.

The term "polypeptide" or "protein" as used herein includes a significant fragment or segment of said protein, and encompasses a stretch of amino acid residues of at least about 8 amino acids, generally at least 10 amino acids, more generally at least 12 amino acids, often at least 14 amino acids, more often at least 16 amino acids, typically at least 18 amino acids, more typically at least 20 amino acids, usually at least 22 amino acids, more usually at least 24 amino acids, preferably at least 26 amino acids, more preferably at least 28 amino acids, and, in particularly preferred embodiments, at least about 30 or more amino acids, e.g., 35, 40, 45, 50, 60, 70, etc.

A "recombinant" nucleic acid is typically defined by its structure. It can be a nucleic acid made by generating a sequence comprising fusion of two fragments which are not naturally contiguous to each other, but is meant to exclude products of nature, e.g., naturally occurring mutant forms.

Certain forms are defined by a method of production. In reference to such, e.g., a product made by a process, the process is use of recombinant nucleic acid techniques, e.g., involving human intervention in the nucleotide sequence, typically selection or production.

Thus, the invention encompasses, for example, nucleic acids comprising sequence derived using a synthetic oligonucleotide process, and products made by transforming cells with a non-naturally occurring vector which encodes these proteins. Such is often done to replace a codon with a redundant codon encoding the same or a conservative amino acid, while typically introducing or removing a sequence recognition site, e.g., for a restriction enzyme. Alternatively, it is performed to

join together nucleic acid segments of desired functions to generate a single genetic entity comprising a desired combination of functions not found in the commonly available natural forms. Restriction enzyme recognition sites are often the target of such artificial manipulations, but other site specific targets, e.g., promoters, DNA replication sites, regulation sequences, control sequences, or other useful features, e.g., primer segments, may be incorporated by design. A similar concept is intended for a recombinant, e.g., fusion, polypeptide. Specifically included are synthetic nucleic acids which, by genetic code redundancy, encode polypeptides similar to fragments of these antigens, and fusions of sequences from various different species variants.

"Solubility" is reflected by sedimentation measured in Svedberg units, which are a measure of the sedimentation velocity of a molecule under particular conditions. The determination of the sedimentation velocity was classically performed in an analytical ultracentrifuge, but is typically now performed in a standard ultracentrifuge. See, Freifelder (1982) Physical Biochemistry (2d ed.) Freeman and Co., San Francisco, CA; and Cantor and Schimmel (1980) Biophysical Chemistry parts 1-3, Freeman and Co., San Francisco, CA. As a crude determination, a sample containing a putatively soluble polypeptide is spun in a standard full sized ultracentrifuge at about 50K rpm for about 10 minutes, and soluble molecules will remain in the supernatant. A soluble particle or polypeptide will typically be less than about 30S, more typically less than about 15S, usually less than about 10S, more usually less than about 6S, and, in particular embodiments, preferably less than about 4S, and more preferably less than about 3S. Solubility of a polypeptide or fragment depends upon the environment and the polypeptide. Many parameters affect polypeptide solubility, including temperature, electrolyte environment, size and molecular characteristics of the polypeptide, and nature of the solvent. Typically, the temperature at which the

polypeptide is used ranges from about 4° C to about 65° C. Usually the temperature at use is greater than about 18° C and more usually greater than about 22° C. For diagnostic purposes, the temperature will usually be about room temperature or warmer, but less than the denaturation temperature of components in the assay. For therapeutic purposes, the temperature will usually be body temperature, typically about 37° C for humans, though under certain situations the temperature may be raised or lowered in situ or in vitro.

The size and structure of the polypeptide should generally be in a substantially stable physiologically active state, and usually not in a denatured state. The polypeptide may be associated with other polypeptides in a quaternary structure, e.g., to confer solubility, or associated with lipids or detergents in a manner which approximates natural lipid bilayer interactions.

The solvent will usually be a biologically compatible buffer, of a type used for preservation of biological activities, and will usually approximate a physiological solvent. Usually the solvent will have a neutral pH, typically between about 5 and 10, and preferably about 7.5. On some occasions, a detergent will be added, typically a mild non-denaturing one, e.g., e.g., CHS (cholesteryl hemisuccinate) or CHAPS (3-([3-cholamidopropyl]dimethyl-ammonio)-1-propane sulfonate), or in a low enough detergent concentration as to avoid significant disruption of structural or physiological properties of the protein.

"Substantially pure" typically means, e.g., in a protein context, that the protein is isolated from other contaminating proteins, nucleic acids, or other biologicals derived from the original source organism. Purity, or "isolation", may be assayed by standard methods, typically by weight, and will ordinarily be at least about 50% pure, more ordinarily at least about 60% pure, generally at least about 70% pure, more generally at least about 80% pure, often at least about 85% pure, more often at least about 90% pure, preferably at least about 95% pure, more preferably at least about 98% pure,

and in most preferred embodiments, at least 99% pure. Carriers or excipients will often be added, or the formulation may be sterile or comprise buffer components.

"Substantial similarity" in the nucleic acid sequence comparison context means either that the segments, or their complementary strands, when compared, are identical when optimally aligned, with appropriate nucleotide insertions or deletions, in at least about 50% of the nucleotides, generally at least 56%, more generally at least 59%, ordinarily at least 62%, more ordinarily at least 65%, often at least 68%, more often at least 71%, typically at least 74%, more typically at least 77%, usually at least 80%, more usually at least about 85%, preferably at least about 90%, more preferably at least about 95 to 98% or more, and in particular embodiments, as high at about 99% or more of the nucleotides. Alternatively, substantial similarity exists when the segments will hybridize under selective hybridization conditions, to a strand, or its complement, typically using a sequence derived from SEQ ID NO: 1, or appropriate parts of 3. Typically, selective hybridization will occur when there is at least about 55% similarity over a stretch of at least about 30 nucleotides, preferably at least about 65% over a stretch of at least about 25 nucleotides, more preferably at least about 75%, and most preferably at least about 90% over about 20 nucleotides. See, Kanehisa (1984) Nucl. Acids Res. 12:203-213. The length of similarity comparison, as described, may be over longer stretches, and in certain embodiments will be over a stretch of at least about 17 nucleotides, usually at least about 20 nucleotides, more usually at least about 24 nucleotides, typically at least about 28 nucleotides, more typically at least about 40 nucleotides, preferably at least about 50 nucleotides, and more preferably at least about 75 to 100 or more nucleotides. The measures of comparison for the SDCMP3 do not reflect on those comparison measures for the SDCMP4.

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are

compared. When using a sequence comparison algorithm, test and reference sequences are input into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. The sequence comparison algorithm then calculates the percent sequence identity for the test sequence(s) relative to the reference sequence, based on the designated program parameters.

Optical alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith and Waterman (1981) Adv. Appl. Math. 2:482, by the homology alignment algorithm of Needleman and Wunsch (1970) J. Mol. Biol. 48:443, by the search for similarity method of Pearson and Lipman (1988) Proc. Nat'l Acad. Sci. USA 85:2444, by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by visual inspection (see generally Ausubel, et al., supra).

One example of a useful algorithm is PILEUP. PILEUP creates a multiple sequence alignment from a group of related sequences using progressive, pairwise alignments to show relationship and percent sequence identity. It also plots a tree or dendrogram showing the clustering relationships used to create the alignment. PILEUP uses a simplification of the progressive alignment method of Feng and Doolittle (1987) J. Mol. Evol. 35:351-360. The method used is similar to the method described by Higgins and Sharp (1989) CABIOS 5:151-153. The program can align up to 300 sequences, each of a maximum length of 5,000 nucleotides or amino acids. The multiple alignment procedure begins with the pairwise alignment of the two most similar sequences, producing a cluster of two aligned sequences. This cluster is then aligned to the next most related sequence or cluster of aligned sequences. Two clusters of sequences are aligned by a simple extension of the pairwise alignment of two individual sequences. The final alignment is achieved by a series of progressive, pairwise alignments. The program is run by designating specific sequences and

their amino acid or nucleotide coordinates for regions of sequence comparison and by designating the program parameters. For example, a reference sequence can be compared to other test sequences to determine the percent sequence identity relationship using the following parameters: default gap weight (3.00), default gap length weight (0.10), and weighted end gaps.

Another example of algorithm that is suitable for determining percent sequence identity and sequence similarity is the BLAST algorithm, which is described Altschul, et al. (1990) J. Mol. Biol. 215:403-410. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul, et al., supra). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLAST program uses as defaults a wordlength (W) of 11, the BLOSUM62 scoring matrix (see Henikoff and Henikoff (1989) Proc. Nat'l Acad. Sci. USA 89:10915) alignments (B) of 50, expectation (E) of 10, M=5, N=4, and a comparison of both strands.

In addition to calculating percent sequence identity, the BLAST algorithm also performs a statistical

analysis of the similarity between two sequences (see, e.g., Karlin and Altschul (1993) Proc. Nat'l Acad. Sci. USA 90:5873-5787). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

A further indication that two nucleic acid sequences of polypeptides are substantially identical is that the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the polypeptide encoded by the second nucleic acid, as described below. Thus, a polypeptide is typically substantially identical to a second polypeptide, for example, where the two peptides differ only by conservative substitutions. Another indication that two nucleic acid sequences are substantially identical is that the two molecules hybridize to each other under stringent conditions, as described below.

"Stringent conditions", in referring to homology or substantial similarity in the hybridization context, will be stringent combined conditions of salt, temperature, organic solvents, and other parameters, typically those controlled in hybridization reactions. The combination of parameters is more important than the measure of any single parameter. See, e.g., Wetmur and Davidson (1968) J. Mol. Biol. 31:349-370. A nucleic acid probe which binds to a target nucleic acid under stringent conditions is specific for said target nucleic acid. Such a probe is typically more than 11 nucleotides in length, and is sufficiently identical or complementary to a target nucleic acid over the region specified by the sequence of the probe to bind the target under stringent hybridization conditions. Generally, a positive signal will exhibit at least 2-fold signal over background,

preferably at least 5-fold, and more preferably at least 15, 25, or even 50 fold over background.

Counterpart SDCMP proteins from other mammalian, e.g., primate or rodent, species can be cloned and isolated by cross-species hybridization of closely related species. See, e.g., below. Similarity may be relatively low between distantly related species, and thus hybridization of relatively closely related species is advisable. Alternatively, preparation of an antibody preparation which exhibits less species specificity may be useful in expression cloning approaches.

The phrase "specifically binds to an antibody" or "specifically immunoreactive with", when referring to a protein or peptide, refers to a binding reaction which is determinative of the presence of the protein in the presence of a heterogeneous population of proteins and other biological components. Thus, under designated immunoassay conditions, the specified antibodies bind to a particular protein and do not significantly bind other proteins present in the sample. Specific binding to an antibody under such conditions may require an antibody that is selected for its specificity for a particular protein. For example, antibodies raised to the human SDCMP3 protein immunogen with the amino acid sequence depicted in SEQ ID NO: 2 can be selected to obtain antibodies specifically immunoreactive with that SDCMP protein and not with other proteins. These antibodies recognize proteins highly similar to the homologous human SDCMP3 protein.

30

III. Nucleic Acids

These SDCMP genes are selectively expressed on dendritic cells. The preferred embodiments, as disclosed, will be useful in standard procedures to isolate genes from other species, e.g., warm blooded animals, such as birds and mammals. Cross hybridization will allow isolation of related proteins from individuals, strains, or species. A number of different approaches are available successfully to isolate a suitable nucleic acid clone based upon the information

40

provided herein. Southern blot hybridization studies should identify homologous genes in other species under appropriate hybridization conditions.

5 Purified protein or defined peptides are useful for generating antibodies by standard methods, as described below. Synthetic peptides or purified protein can be presented to an immune system to generate polyclonal and monoclonal antibodies. See, e.g., Coligan (1991) Current Protocols in Immunology Wiley/Greene, NY; and Harlow and
10 Lane (1989) Antibodies: A Laboratory Manual Cold Spring Harbor Press, NY, which are incorporated herein by reference. Alternatively, a SDCMP antigen binding composition can be useful as a specific binding reagent, and advantage can be taken of its specificity of binding,
15 for, e.g., purification of a SDCMP protein.

The specific binding composition can be used for screening an expression library made from a cell line which expresses the respective SDCMP protein. Many methods for screening are available, e.g., standard
20 staining of surface expressed ligand, or by panning. Screening of intracellular expression can also be performed by various staining or immunofluorescence procedures. The binding compositions could be used to affinity purify or sort out cells expressing the antigen.

25 Sequence analysis suggests these SDCMPs are members of the lectin/asialoglycoprotein superfamily of receptors. See also USSN 60/053,080, which is incorporated herein by reference.

Sequences encoding a primate SDCMP3, initially
30 designated lectin73, isolated from a dendritic cell library, are shown in SEQ ID NO: 1 and 2. An ORF runs from about 108 to 593. Comparison with rodent SDCMP3 suggests that the sequence may be truncated at or near the C-terminus. Sequence encoding the mouse counterpart
35 are shown in SEQ ID NO: 3 and 4.

Analysis of the human SDCMP3 suggests that the protein is a type II membrane protein, with the transmembrane segment running from about ser22 to thr42. The cytoplasmic tail would be at the N terminus, from
40 met1 to trp21. A C-type lectin domain corresponds to

16

about cys79 to arg162. The human protein has a predicted molecular weight of about 18,500 daltons, with an isoelectric point of about 6, and a charge of about -2.6 at pH 7. Hydrophilicity analysis indicates significant stretches of hydrophilic sequence from about 1-22, 42-63, 94-106, and 142-162. Such segments will likely be more antigenic. Similar analysis of the mouse SDCMP3 suggests that the protein is also a type II membrane protein with the transmembrane segment running from about ser20 to thr40. The cytoplasmic tail would then run from about met1 to trp19; and the C-type lectin domain would correspond to about cys79 to at least arg162. Two putative N-glycosylation sites correspond to asn131-ser133 and asn183-ser185. Computationally identified particularly antigenic stretches for the human will run from about met1-ser18; tyr43-arg53; lys72-ser85; ser94-asn106; and ser135-arg162. See, e.g., Beattie, et al. (1992) Eur. J. Biochem. 210:59-66. Alignment of primate SDCMP3 with rodent SDCMP3 is shown in the following table:

TABLE

pr	MMQEQQPQST	EKRGWLSLRL	WSVAGISIAL	LSACFIVSCV	VTYHFTYGET
ro	MVQERQSQGK	-GVCWT-LRL	WSAAVISMML	LSTCFIASCV	VTYQFIMDQP
pr	GKRLSELHSY	HSSLTCFSEG	TKV--PAWGC	CPASWKSFGS	SCYFISSEEK
ro	SRRLYELHTY	HSSLTCFSEG	TMVSEKMWGC	CPNHWSFGS	SCYLISTKEN
pr	VWSKSEQNCV	EMGAHLVVFN	TEAEQNFIVQ	QLNESFSYFL	GLSDPQGNNN
ro	FWSTSEQNCV	QMG AHLVVIN	TEAEQNFITQ	QLNESLSYFL	GLSDPQGNNGK
pr	WQWIDKTPYE	KNVR			
ro	WQWIDDTFFS	QNVRFWHPHE	PNLPEERCVS	IVYWNPSKWG	WNDVFCDSKH
ro	NSICEMKKIY	L			

Analysis of the human SDCMP4 suggests that the protein is a type II membrane protein. There are two forms, the long form, initially designated lectin 47, isolated from a dendritic cell library (SEQ ID NO: 5 and 6), and the short form (SEQ ID NO: 7 and 8), which corresponds to a deletion of nucleotides 362-499 of the long form, and which may result from an alternative splice event. Differences also exist at nucleotides 108

and 239. Assorted variations in sequence may reflect sequencing errors, or allelic variants.

5 The predicted transmembrane segment of the long form runs from about leu45 to met67. The amino proximal portion of the protein would be cytoplasmic. Computationally identified particularly antigenic stretches for the human will run from about met1-arg44; trp70-thr113; and asn139-cys220. A notable feature is the internalization motif (GYTQ, residues 14-17) into
10 intracytoplasmic domain. The CRD would extend from about cys120 to met247 of the long form, and from about cys74 to met201 of the short form. The long form would be predicted to have a molecular weight of about 27.6 kD, and the short form about 22.5 kD with a calculated
15 isoelectric point of about 4.6, and a charge of -7.8 at pH 7.

The extracellular domain of the SDCMP4 proteins contain a C-type (Ca++ dependent) lectin carbohydrate recognition domain (CRD), as indicated by significant
20 sequence homology with other lectins. The prototype of the type II transmembrane C-type lectins is the hepatic asialoglycoprotein-receptor (ASGPR).

The CRD of the hepatic ASGPR displays binding specificity for galactose. In addition, the
25 intracellular domain of the ASGPR bears a tyrosine-based motif that enables ligand internalization. Unlike the ASGPR or the macrophage mannose-receptor, the CRD sequence of SDCMP4 does not as strongly suggest its sugar specificity. Such lack of suggestion is also a feature
30 of other C-type lectins, as exemplified by the NGK2 receptors on NK cells.

The intracellular domains of both embodiments SDCMP4 display an internalization sequence (YTQL) of the YXXØ type, where Ø represents a hydrophobic amino acid. As
35 reference, the internalization motif of the liver ASGPR-H1 chain is YQDL.

Notably, several type II transmembrane C-type lectins (e.g., human NKG2 and DC-IR, mouse Ly49 and NKRP1) are members of the immunoreceptor superfamily (IRS) system. Some forms of these receptors have the
40

ability to deliver an inhibitory signal through an intracellular ITIM motif. By contrast, other forms lack an ITIM motif, and as such do not transmit a negative signal. A hallmark of such non-inhibitory IRS members is the presence of a charged amino-acid in the transmembrane region. Alternatively, truncated forms may interact with transmembrane accessory molecules. See, e.g., Lanier, et al. (1998) Nature 391:703-7; and USSN 60/069,639, which are both incorporated herein by reference.

SDCMP4 neither displays an ITIM motif in its intracellular domain, nor a charged transmembrane residue. On this basis, it appears unlikely that SDCMP4 defines a new family of C-type lectin IRS genes. Rather, it can be suggested that SDCMP4 is related to the ASGPR system of molecules involved in ligand internalization.

Two forms of SDCMP4 have been identified, that differ by the presence of a 46 amino-acid membrane-proximal insertion in the extracellular domain. Insertions in this region also occur in the macrophage and the dendritic cell (ETA10) ASGPRs.

Finally, expression of SDCMP4 has been observed by RT-PCR in myeloid cells (dendritic cells, monocytes, and granulocytes). In contrast to SDCMP3, expression of SDCMP4 is not down-regulated in DC following activation by PMA and ionomycin.

The CRD of the hepatic ASGPR displays binding specificity for galactose. In addition, the intracellular domain of the ASGPR bears a tyrosine-based motif that enables ligand internalization. Unlike the hepatic ASGPR or the macrophage mannose-receptor, the CRD sequence of SDCMP4 does not as strongly suggest its sugar specificity. Such lack of suggestion is also a feature of other C-type lectins, as exemplified by the NGK2 receptors on NK cells.

The intracellular domain of SDCMP4 displays an internalization sequence (YTQL) of the YXXØ type, where Ø represents a hydrophobic amino acid. As reference, the internalization motif of the liver ASGPR-H1 chain is YQDL.

Notably, several type II transmembrane C-type lectins (e.g., human NKG2 and DC-IR, mouse Ly49 and NKRP1) are members of the immunoreceptor superfamily (IRS) system. Some forms of these receptors have the ability to deliver an inhibitory signal through an intracellular ITIM motif. By contrast, other forms lack an ITIM motif, and as such do not transmit a negative signal. A hallmark of such non-inhibitory IRS members is the presence of a charged amino-acid in the transmembrane region.

SDCMP4 neither displays an ITIM motif in its intracellular domain, nor a charged transmembrane residue. On this basis, it appears unlikely that SDCMP4 defines a new family of C-type lectin IRS genes. Rather, it can be suggested that SDCMP4 is related to the ASGPR system of molecules involved in ligand internalization.

Two forms of SDCMP4 have been identified, that differ by the presence of a 46 amino-acid membrane-proximal insertion in the extracellular domain. Insertions in this region also occur in the macrophage and the dendritic cell (ETA10) ASGPRs.

Finally, expression of SDCMP4 has been observed by RT-PCR in myeloid cells (dendritic cells, monocytes, and granulocytes). In contrast to SDCMP3, expression of SDCMP4 is not down-regulated in DC following activation by PMA and ionomycin.

Close sequences to these are the ETA10 sequences. See, e.g., Suzuki, et al. (1996) J. Immunol. 156:128-135; and Sato, et al. (1992) J. Biochem. 111:331-336. The extracellular domain displays a number of features indicative of a C-type (Ca⁺⁺ dependent) carbohydrate recognition domains (CRD). While the CRD of the human form appears truncated at its carboxyl terminus, the CRD of the mouse homolog (1469D4) is not truncated and clearly classifies the lectin as a novel member of the C-type superfamily.

The prototype of the C-type transmembrane type II lectins is the hepatic asialoglycoprotein-receptor (ASGPR). The ASGPR, however, bears an intracytoplasmic tyrosine-based ligand internalization sequence, that is

found neither in the human or mouse SDCMP3. The gene encoding human SDCMP3 maps on chromosome 12 p12-13, e.g., in the human NK receptor complex. Notably, this region includes the NKG2 genes and the CD94 gene, which encode C-type transmembrane type II lectins and represent examples of the immunoreceptor superfamily (IRS) system. Thus, killer-cell inhibitory receptors (KIR) CD94-NKG2A/B heterodimers transduce a negative signal by virtue of an intracellular tyrosine-based ITIM motif in the NKG2 sequences. However, the other forms of NKG2 lack an ITIM motif, and the heterodimers resulting with CD94 are non-inhibitory.

The intracellular domain of human SDCMP3 does not contain an ITIM motif. However, on the basis of its chromosomal localization, as well as its significant (36.2%) homology with the IRS gene DC-IR, it is predicted to be a member of a novel C-type lectin family of IRS genes. By analogy with other IRS genes, it is likely that the SDCMP3 represents a family of genes which will comprise several members, either with inhibitory (ITIM) or non-inhibitory function.

By RT-PCR, primate SDCMP3 expression is restricted to myeloid cells, being observed in dendritic cells (DC), monocytes, and macrophages. Expression is selectively seen in CD14-derived DC, rather than in CD1a-derived Langerhans-type DC. Finally, expression of SDCMP3 is downregulated by activation with PMA with ionomycin.

The peptide segments can also be used to design and produce appropriate oligonucleotides to screen a library to determine the presence of a similar gene, e.g., an identical or polymorphic variant, or to identify a DC. The genetic code can be used to select appropriate oligonucleotides useful as probes for screening. In combination with polymerase chain reaction (PCR) techniques, synthetic oligonucleotides will be useful in selecting desired clones from a library.

Complementary sequences will also be used as probes or primers. Based upon identification of the likely amino terminus, other peptides should be particularly useful, e.g., coupled with anchored vector or poly-A

complementary PCR techniques or with complementary DNA of other peptides.

Techniques for nucleic acid manipulation of genes encoding these DC proteins, e.g., subcloning nucleic acid sequences encoding polypeptides into expression vectors, labeling probes, DNA hybridization, and the like are described generally in Sambrook, et al. (1989) Molecular Cloning: A Laboratory Manual (2nd ed.) Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor Press, NY, which is incorporated herein by reference and hereinafter referred to as "Sambrook, et al." See also, Coligan, et al. (1987 and periodic supplements) Current Protocols in Molecular Biology Greene/Wiley, New York, NY, referred to as "Coligan, et al."

There are various methods of isolating the DNA sequences encoding these DC proteins. For example, DNA is isolated from a genomic or cDNA library using labeled oligonucleotide probes having sequences identical or complementary to the sequences disclosed herein. Full-length probes may be used, or oligonucleotide probes may be generated by comparison of the sequences disclosed with other proteins and selecting specific primers. Such probes can be used directly in hybridization assays to isolate DNA encoding DC proteins, or probes can be designed for use in amplification techniques such as PCR, for the isolation of DNA encoding DC proteins.

To prepare a cDNA library, mRNA is isolated from cells which express the DC protein. cDNA is prepared from the mRNA and ligated into a recombinant vector. The vector is transfected into a recombinant host for propagation, screening and cloning. Methods for making and screening cDNA libraries are well known. See Gubler and Hoffman (1983) Gene 25:263-269; Sambrook, et al.; or Coligan, et al.

For a genomic library, the DNA can be extracted from tissue and either mechanically sheared or enzymatically digested to yield fragments of about 12-20 kb. The fragments are then separated by gradient centrifugation and cloned in bacteriophage lambda vectors. These vectors and phage are packaged in vitro, as described,

e.g., in Sambrook, et al. or Coligan, et al. Recombinant phage are analyzed by plaque hybridization as described in Benton and Davis (1977) Science 196:180-182. Colony hybridization is carried out as generally described in, e.g., Grunstein, et al. (1975) Proc. Natl. Acad. Sci. USA 72:3961-3965.

DNA encoding a DC protein can be identified in either cDNA or genomic libraries by its ability to hybridize with the nucleic acid probes described herein, for example in colony or plaque hybridization experiments. The corresponding DNA regions are isolated by standard methods familiar to those of skill in the art. See Sambrook, et al.

Various methods of amplifying target sequences, such as the polymerase chain reaction, can also be used to prepare DNA encoding DC proteins. Polymerase chain reaction (PCR) technology is used to amplify such nucleic acid sequences directly from mRNA, from cDNA, and from genomic libraries or cDNA libraries. The isolated sequences encoding DC proteins may also be used as templates for PCR amplification.

In PCR techniques, oligonucleotide primers complementary to two 5' regions in the DNA region to be amplified are synthesized. The polymerase chain reaction is then carried out using the two primers. See Innis, et al. (eds.) (1990) PCR Protocols: A Guide to Methods and Applications Academic Press, San Diego, CA. Primers can be selected to amplify the entire regions encoding a selected full-length DC protein or to amplify smaller DNA segments as desired. In particular, the provided sequences provide primers of, e.g., 15-30 nucleotides, which can be used to amplify the desired coding sequences, or fragments thereof. Once such regions are PCR-amplified, they can be sequenced and oligonucleotide probes can be prepared from sequence obtained using standard techniques. These probes can then be used to isolate DNAs encoding other forms of the DC proteins.

Oligonucleotides for use as probes are chemically synthesized according to the solid phase phosphoramidite triester method first described by Beaucage and

Carruthers (1983) Tetrahedron Lett. 22(20):1859-1862, or using an automated synthesizer, as described in Needham-VanDevanter, et al. (1984) Nucleic Acids Res. 12:6159-6168. Purification of oligonucleotides is performed, e.g., by native acrylamide gel electrophoresis or by anion-exchange HPLC as described in Pearson and Regnier (1983) J. Chrom. 255:137-149. The sequence of the synthetic oligonucleotide can be verified using the chemical degradation method of Maxam and Gilbert in Grossman and Moldave (eds. 1980) Methods in Enzymology 65:499-560 Academic Press, New York.

This invention provides isolated DNA or fragments to encode a DC protein, as described. In addition, this invention provides isolated or recombinant DNA which encodes a biologically active protein or polypeptide which is capable of hybridizing under appropriate conditions, e.g., high stringency, with the DNA sequences described herein. Said biologically active protein or polypeptide can be a naturally occurring form, or a recombinant protein or fragment, and have an amino acid sequence as disclosed in SEQ ID NO: 2, 4, 6, or 8. Preferred embodiments will be full length natural isolates, e.g., from a primate. In glycosylated form, the proteins should exhibit larger sizes. Further, this invention encompasses the use of isolated or recombinant DNA, or fragments thereof, which encode proteins which are homologous to each respective DC protein. The isolated DNA can have the respective regulatory sequences in the 5' and 3' flanks, e.g., promoters, enhancers, poly-A addition signals, and others.

IV. Making DC Gene Products

DNAs which encode these DC proteins or fragments thereof can be obtained by chemical synthesis, screening cDNA libraries, or by screening genomic libraries prepared from a wide variety of cell lines or tissue samples.

These DNAs can be expressed in a wide variety of host cells for the synthesis of a full-length protein or

fragments which can, e.g., be used to generate polyclonal or monoclonal antibodies; for binding studies; for construction and expression of modified molecules; and for structure/function studies. Each of these DC proteins or their fragments can be expressed in host cells that are transformed or transfected with appropriate expression vectors. These molecules can be substantially purified to be free of protein or cellular contaminants, other than those derived from the recombinant host, and therefore are particularly useful in pharmaceutical compositions when combined with a pharmaceutically acceptable carrier and/or diluent. The antigen, or portions thereof, may be expressed as fusions with other proteins.

Expression vectors are typically self-replicating DNA or RNA constructs containing the desired DC gene or its fragments, usually operably linked to suitable genetic control elements that are recognized in a suitable host cell. These control elements are capable of effecting expression within a suitable host. The specific type of control elements necessary to effect expression will depend upon the eventual host cell used. Generally, the genetic control elements can include a prokaryotic promoter system or a eukaryotic promoter expression control system, and typically include a transcriptional promoter, an optional operator to control the onset of transcription, transcription enhancers to elevate the level of mRNA expression, a sequence that encodes a suitable ribosome binding site, and sequences that terminate transcription and translation. Expression vectors also usually contain an origin of replication that allows the vector to replicate independently from the host cell.

The vectors of this invention contain DNAs which encode the various DC proteins, or a fragment thereof, typically encoding, e.g., a biologically active polypeptide, or protein. The DNA can be under the control of a viral promoter and can encode a selection marker. This invention further contemplates use of such expression vectors which are capable of expressing

eukaryotic cDNA coding for a DC protein in a prokaryotic or eukaryotic host, where the vector is compatible with the host and where the eukaryotic cDNA coding for the protein is inserted into the vector such that growth of the host containing the vector expresses the cDNA in question. Usually, expression vectors are designed for stable replication in their host cells or for amplification to greatly increase the total number of copies of the desirable gene per cell. It is not always necessary to require that an expression vector replicate in a host cell, e.g., it is possible to effect transient expression of the protein or its fragments in various hosts using vectors that do not contain a replication origin that is recognized by the host cell. It is also possible to use vectors that cause integration of a DC gene or its fragments into the host DNA by recombination, or to integrate a promoter which controls expression of an endogenous gene.

Vectors, as used herein, comprise plasmids, viruses, bacteriophage, integratable DNA fragments, and other vehicles which enable the integration of DNA fragments into the genome of the host. Expression vectors are specialized vectors which contain genetic control elements that effect expression of operably linked genes. Plasmids are the most commonly used form of vector but all other forms of vectors which serve an equivalent function are suitable for use herein. See, e.g., Pouwels, et al. (1985 and Supplements) Cloning Vectors: A Laboratory Manual Elsevier, N.Y.; and Rodriguez, et al. (eds. 1988) Vectors: A Survey of Molecular Cloning Vectors and Their Uses Butterworth, Boston, MA.

Suitable host cells include prokaryotes, lower eukaryotes, and higher eukaryotes. Prokaryotes include both gram negative and gram positive organisms, e.g., *E. coli* and *B. subtilis*. Lower eukaryotes include yeasts, e.g., *S. cerevisiae* and *Pichia*, and species of the genus *Dictyostelium*. Higher eukaryotes include established tissue culture cell lines from animal cells, both of non-mammalian origin, e.g., insect cells, and birds, and of mammalian origin, e.g., human, primates, and rodents.

Prokaryotic host-vector systems include a wide variety of vectors for many different species. As used herein, *E. coli* and its vectors will be used generically to include equivalent vectors used in other prokaryotes.

5 A representative vector for amplifying DNA is pBR322 or its derivatives. Vectors that can be used to express DC proteins or fragments include, but are not limited to, such vectors as those containing the lac promoter (pUC-series); trp promoter (pBR322-trp); Ipp promoter (the

10 pIN-series); lambda-pP or pR promoters (pOTS); or hybrid promoters such as ptac (pDR540). See Brosius, et al. (1988) "Expression Vectors Employing Lambda-, trp-, lac-, and Ipp-derived Promoters", in Rodriguez and Denhardt (eds.) Vectors: A Survey of Molecular Cloning Vectors and Their Uses 10:205-236 Buttersworth, Boston, MA.

15 Lower eukaryotes, e.g., yeasts and *Dictyostelium*, may be transformed with DC gene sequence containing vectors. For purposes of this invention, the most common lower eukaryotic host is the baker's yeast, *Saccharomyces cerevisiae*. It will be used generically to represent

20 lower eukaryotes although a number of other strains and species are also available. Yeast vectors typically consist of a replication origin (unless of the integrating type), a selection gene, a promoter, DNA encoding the desired protein or its fragments, and

25 sequences for translation termination, polyadenylation, and transcription termination. Suitable expression vectors for yeast include such constitutive promoters as 3-phosphoglycerate kinase and various other glycolytic enzyme gene promoters or such inducible promoters as the

30 alcohol dehydrogenase 2 promoter or metallothionein promoter. Suitable vectors include derivatives of the following types: self-replicating low copy number (such as the YRp-series), self-replicating high copy number

35 (such as the YEpl-series); integrating types (such as the YIp-series), or mini-chromosomes (such as the YCp-series).

Higher eukaryotic tissue culture cells are the preferred host cells for expression of the DC protein.

40 In principle, most any higher eukaryotic tissue culture

cell line may be used, e.g., insect baculovirus expression systems, whether from an invertebrate or vertebrate source. However, mammalian cells are preferred to achieve proper processing, both
5 cotranslationally and posttranslationally. Transformation or transfection and propagation of such cells is routine. Useful cell lines include HeLa cells, Chinese hamster ovary (CHO) cell lines, baby rat kidney (BRK) cell lines, insect cell lines, bird cell lines, and monkey (COS) cell
10 lines. Expression vectors for such cell lines usually include an origin of replication, a promoter, a translation initiation site, RNA splice sites (e.g., if genomic DNA is used), a polyadenylation site, and a transcription termination site. These vectors also may
15 contain a selection gene or amplification gene. Suitable expression vectors may be plasmids, viruses, or retroviruses carrying promoters derived, e.g., from such sources as from adenovirus, SV40, parvoviruses, vaccinia virus, or cytomegalovirus. Representative examples of
20 suitable expression vectors include pCDNA1; pCD, see Okayama, et al. (1985) Mol. Cell Biol. 5:1136-1142; pMC1neo Poly-A, see Thomas, et al. (1987) Cell 51:503-512; and a baculovirus vector such as pAC 373 or pAC 610.

In certain instances, the DC proteins need not be
25 glycosylated to elicit biological responses in certain assays. However, it will often be desirable to express a DC polypeptide in a system which provides a specific or defined glycosylation pattern. In this case, the usual pattern will be that provided naturally by the expression
30 system. However, the pattern will be modifiable by exposing the polypeptide, e.g., in unglycosylated form, to appropriate glycosylating proteins introduced into a heterologous expression system. For example, a DC gene may be co-transformed with one or more genes encoding
35 mammalian or other glycosylating enzymes. It is further understood that over glycosylation may be detrimental to DC protein biological activity, and that one of skill may perform routine testing to optimize the degree of glycosylation which confers optimal biological activity.

A DC protein, or a fragment thereof, may be engineered to be phosphatidyl inositol (PI) linked to a cell membrane, but can be removed from membranes by treatment with a phosphatidyl inositol cleaving enzyme, e.g., phosphatidyl inositol phospholipase-C. This releases the antigen in a biologically active form, and allows purification by standard procedures of protein chemistry. See, e.g., Low (1989) Biochem. Biophys. Acta 988:427-454; Tse, et al. (1985) Science 230:1003-1008; Brunner, et al. (1991) J. Cell Biol. 114:1275-1283; and Coligan, et al. (eds.) (1996 and periodic supplements) Current Protocols in Protein Science, John Wiley and Sons, New York, NY.

Now that these SDCMP proteins have been characterized, fragments or derivatives thereof can be prepared by conventional processes for synthesizing peptides. These include processes such as are described in Stewart and Young (1984) Solid Phase Peptide Synthesis Pierce Chemical Co., Rockford, IL; Bodanszky and Bodanszky (1984) The Practice of Peptide Synthesis Springer-Verlag, New York, NY; and Bodanszky (1984) The Principles of Peptide Synthesis Springer-Verlag, New York, NY. See also Merrifield (1986) Science 232:341-347; and Dawson, et al. (1994) Science 266:776-779. For example, an azide process, an acid chloride process, an acid anhydride process, a mixed anhydride process, an active ester process (for example, p-nitrophenyl ester, N-hydroxysuccinimide ester, or cyanomethyl ester), a carbodiimidazole process, an oxidative-reductive process, or a dicyclohexylcarbodiimide (DCCD)/additive process can be used. Solid phase and solution phase syntheses are both applicable to the foregoing processes.

The prepared protein and fragments thereof can be isolated and purified from the reaction mixture by means of peptide separation, for example, by extraction, precipitation, electrophoresis and various forms of chromatography, and the like. The DC proteins of this invention can be obtained in varying degrees of purity depending upon the desired use. Purification can be accomplished by use of known protein purification

techniques or by the use of the antibodies or binding partners herein described, e.g., in immunoabsorbant affinity chromatography. This immunoabsorbant affinity chromatography is carried out by first linking the antibodies to a solid support and contacting the linked antibodies with solubilized lysates of appropriate source cells, lysates of other cells expressing the protein, or lysates or supernatants of cells producing the proteins as a result of DNA techniques, see below.

Multiple cell lines may be screened for one which expresses said protein at a high level compared with other cells. Various cell lines, e.g., a mouse thymic stromal cell line TA4, is screened and selected for its favorable handling properties. Natural DC cell proteins can be isolated from natural sources, or by expression from a transformed cell using an appropriate expression vector. Purification of the expressed protein is achieved by standard procedures, or may be combined with engineered means for effective purification at high efficiency from cell lysates or supernatants. FLAG or His₆ segments can be used for such purification features.

V. Antibodies

Antibodies can be raised to the various DC proteins, including individual, polymorphic, allelic, strain, or species variants, and fragments thereof, both in their naturally occurring (full-length) forms and in their recombinant forms. Additionally, antibodies can be raised to DC proteins in either their active forms or in their inactive forms. Anti-idiotypic antibodies may also be used.

a. Antibody Production

A number of immunogens may be used to produce antibodies specifically reactive with these DC proteins. Recombinant protein is the preferred immunogen for the production of monoclonal or polyclonal antibodies. Naturally occurring protein may also be used either in pure or impure form. Synthetic peptides made using the human DC protein sequences described herein may also be used as an immunogen for the production of antibodies to the

DC protein. Recombinant protein can be expressed in eukaryotic or prokaryotic cells as described herein, and purified as described. The product is then injected into an animal capable of producing antibodies. Either
5 monoclonal or polyclonal antibodies may be generated for subsequent use in immunoassays to measure the protein.

Methods of producing polyclonal antibodies are known to those of skill in the art. In brief, an immunogen, preferably a purified protein, is mixed with an adjuvant and animals are immunized with the mixture. The animal's
10 immune response to the immunogen preparation is monitored by taking test bleeds and determining the titer of reactivity to the DC protein of interest. When appropriately high titers of antibody to the immunogen
15 are obtained, blood is collected from the animal and antisera are prepared. Further fractionation of the antisera to enrich for antibodies reactive to the protein can be done if desired. See, e.g., Harlow and Lane.

Monoclonal antibodies may be obtained by various
20 techniques familiar to those skilled in the art. Briefly, spleen cells from an animal immunized with a desired antigen are immortalized, commonly by fusion with a myeloma cell. See, e.g., Kohler and Milstein (1976) Eur. J. Immunol. 6:511-519, which is incorporated herein
25 by reference. Alternative methods of immortalization include transformation with Epstein Barr Virus, oncogenes, or retroviruses, or other methods known in the art. Colonies arising from single immortalized cells are screened for production of antibodies of the desired
30 specificity and affinity for the antigen, and yield of the monoclonal antibodies produced by such cells may be enhanced by various techniques, including injection into the peritoneal cavity of a vertebrate host.

Alternatively, one may isolate DNA sequences which encode
35 a monoclonal antibody or a binding fragment thereof by screening a DNA library from human B cells according to the general protocol outlined by Huse, et al. (1989) Science 246:1275-1281.

Antibodies, including binding fragments and single
40 chain versions, against predetermined fragments of these

DC proteins can be raised by immunization of animals with conjugates of the fragments with carrier proteins as described above. Monoclonal antibodies are prepared from cells secreting the desired antibody. These antibodies
5 can be screened for binding to normal or defective DC proteins, or screened for agonistic or antagonistic activity. These monoclonal antibodies will usually bind with at least a K_D of about 1 mM, more usually at least about 300 μ M, typically at least about 10 μ M, more
10 typically at least about 30 μ M, preferably at least about 10 μ M, and more preferably at least about 3 μ M or better.

In some instances, it is desirable to prepare monoclonal antibodies from various mammalian hosts, such as mice, rodents, primates, humans, etc. Description of
15 techniques for preparing such monoclonal antibodies may be found in, e.g., Stites, et al. (eds.) Basic and Clinical Immunology (4th ed.) Lange Medical Publications, Los Altos, CA, and references cited therein; Harlow and Lane (1988) Antibodies: A Laboratory Manual CSH Press;
20 Goding (1986) Monoclonal Antibodies: Principles and Practice (2d ed.) Academic Press, New York, NY; and particularly in Kohler and Milstein (1975) Nature 256:495-497, which discusses one method of generating monoclonal antibodies. Summarized briefly, this method
25 involves injecting an animal with an immunogen to initiate a humoral immune response. The animal is then sacrificed and cells taken from its spleen, which are then fused with myeloma cells. The result is a hybrid cell or "hybridoma" that is capable of reproducing in
30 vitro. The population of hybridomas is then screened to isolate individual clones, each of which secretes a single antibody species to the immunogen. In this manner, the individual antibody species obtained are the products of immortalized and cloned single B cells from
35 the immune animal generated in response to a specific site recognized on the immunogenic substance.

Other suitable techniques involve selection of libraries of antibodies in phage or similar vectors. See, Huse, et al. (1989) "Generation of a Large
40 Combinatorial Library of the Immunoglobulin Repertoire in

Phage Lambda," Science 246:1275-1281; and Ward, et al. (1989) Nature 341:544-546. The polypeptides and antibodies of the present invention may be used with or without modification, including chimeric or humanized antibodies. Frequently, the polypeptides and antibodies will be labeled by joining, either covalently or non-covalently, a substance which provides for a detectable signal. A wide variety of labels and conjugation techniques are known and are reported extensively in both the scientific and patent literature. Suitable labels include radionuclides, enzymes, substrates, cofactors, inhibitors, fluorescent moieties, chemiluminescent moieties, magnetic particles, and the like. Patents, teaching the use of such labels include U.S. Patent Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241. Also, recombinant immunoglobulins may be produced. See, Cabilly, U.S. Patent No. 4,816,567; and Queen, et al. (1989) Proc. Nat'l Acad. Sci. USA 86:10029-10033.

The antibodies of this invention can also be used for affinity chromatography in isolating each DC protein. Columns can be prepared where the antibodies are linked to a solid support, e.g., particles, such as agarose, SEPHADEX, or the like, where a cell lysate may be passed through the column, the column washed, followed by increasing concentrations of a mild denaturant, whereby purified DC protein will be released.

The antibodies may also be used to screen expression libraries for particular expression products. Usually the antibodies used in such a procedure will be labeled with a moiety allowing easy detection of presence of antigen by antibody binding.

Antibodies to SDCMP proteins may be used for the analysis or, or identification of specific cell population components which express the respective protein. By assaying the expression products of cells expressing DC proteins it is possible to diagnose disease, e.g., immune-compromised conditions, DC depleted conditions, or overproduction of DC.

Antibodies raised against each DC will also be useful to raise anti-idiotypic antibodies. These will be useful in detecting or diagnosing various immunological conditions related to expression of the respective antigens.

b. Immunoassays

A particular protein can be measured by a variety of immunoassay methods. For a review of immunological and immunoassay procedures in general, see Stites and Terr (eds.) 1991 Basic and Clinical Immunology (7th ed.). Moreover, the immunoassays of the present invention can be performed in any of several configurations, which are reviewed extensively in Maggio (ed. 1980) Enzyme Immunoassay CRC Press, Boca Raton, Florida; Tijan (1985) "Practice and Theory of Enzyme Immunoassays," Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers B.V., Amsterdam; and Harlow and Lane Antibodies, A Laboratory Manual, supra, each of which is incorporated herein by reference. See also Chan (ed.) (1987) Immunoassay: A Practical Guide Academic Press, Orlando, FL; Price and Newman (eds.) (1991) Principles and Practice of Immunoassays Stockton Press, NY; and Ngo (ed. 1988) Non-isotopic Immunoassays Plenum Press, NY.

Immunoassays for measurement of these DC proteins can be performed by a variety of methods known to those skilled in the art. In brief, immunoassays to measure the protein can be competitive or noncompetitive binding assays. In competitive binding assays, the sample to be analyzed competes with a labeled analyte for specific binding sites on a capture agent bound to a solid surface. Preferably the capture agent is an antibody specifically reactive with the DC protein produced as described above. The concentration of labeled analyte bound to the capture agent is inversely proportional to the amount of free analyte present in the sample.

In a competitive binding immunoassay, the DC protein present in the sample competes with labeled protein for binding to a specific binding agent, for example, an antibody specifically reactive with the DC protein. The

binding agent may be bound to a solid surface to effect separation of bound labeled protein from the unbound labeled protein. Alternately, the competitive binding assay may be conducted in liquid phase and any of a variety of techniques known in the art may be used to separate the bound labeled protein from the unbound labeled protein. Following separation, the amount of bound labeled protein is determined. The amount of protein present in the sample is inversely proportional to the amount of labeled protein binding.

Alternatively, a homogenous immunoassay may be performed in which a separation step is not needed. In these immunoassays, the label on the protein is altered by the binding of the protein to its specific binding agent. This alteration in the labeled protein results in a decrease or increase in the signal emitted by label, so that measurement of the label at the end of the immunoassay allows for detection or quantitation of the protein.

These DC proteins may also be quantitatively determined by a variety of noncompetitive immunoassay methods. For example, a two-site, solid phase sandwich immunoassay may be used. In this type of assay, a binding agent for the protein, for example an antibody, is attached to a solid support. A second protein binding agent, which may also be an antibody, and which binds the protein at a different site, is labeled. After binding at both sites on the protein has occurred, the unbound labeled binding agent is removed and the amount of labeled binding agent bound to the solid phase is measured. The amount of labeled binding agent bound is directly proportional to the amount of protein in the sample.

Western blot analysis can be used to determine the presence of DC proteins in a sample. Electrophoresis is carried out, e.g., on a tissue sample suspected of containing the protein. Following electrophoresis to separate the proteins, and transfer of the proteins to a suitable solid support such as a nitrocellulose filter, the solid support is incubated with an antibody reactive

with the denatured protein. This antibody may be labeled, or alternatively may be it may be detected by subsequent incubation with a second labeled antibody that binds the primary antibody.

5 The immunoassay formats described above employ labeled assay components. The label can be in a variety of forms. The label may be coupled directly or indirectly to the desired component of the assay according to methods well known in the art. A wide
10 variety of labels may be used. The component may be labeled by any one of several methods. Traditionally a radioactive label incorporating ^3H , ^{125}I , ^{35}S , ^{14}C , or ^{32}P is used. Non-radioactive labels include ligands
15 which bind to labeled antibodies, fluorophores, chemiluminescent agents, enzymes, and antibodies which can serve as specific binding pair members for a labeled protein. The choice of label depends on sensitivity required, ease of conjugation with the compound, stability requirements, and available instrumentation.
20 For a review of various labeling or signal producing systems which may be used, see U.S. Patent No. 4,391,904, which is incorporated herein by reference.

 Antibodies reactive with a particular protein can also be measured by a variety of immunoassay methods.
25 For reviews of immunological and immunoassay procedures applicable to the measurement of antibodies by immunoassay techniques, see, e.g., Stites and Terr (eds.) Basic and Clinical Immunology (7th ed.) supra; Maggio (ed.) Enzyme Immunoassay, supra; and Harlow and Lane
30 Antibodies, A Laboratory Manual, supra.

 A variety of different immunoassay formats, separation techniques, and labels can be also be used similar to those described above for the measurement of specific proteins.

35

VI. Purified SDCMP proteins

 Primate, e.g., human, SDCMP3 nucleotide and amino acid sequences are provided in SEQ ID NO: 1 and 2; rodent, e.g., mouse SDCMP3 sequences are provided in SEQ
40 ID NO: 3 and 4. Primate, e.g., human SDCMP4 nucleotide

and amino acid sequences are provided in SEQ ID NO: 5, 6, 7, and 8. The peptide sequences allow preparation of peptides to generate antibodies to recognize such segments, and allow preparation of oligonucleotides which encode such sequences.

Standard methods of purification are available, and the purification may be followed by use of specific antibodies.

VII. Physical Variants

This invention also encompasses proteins or peptides having substantial amino acid sequence similarity with an amino acid sequence of a SEQ ID NO: 2, 4, 6, or 8. Variants exhibiting substitutions, e.g., 20 or fewer, preferably 10 or fewer, and more preferably 5 or fewer substitutions, are also enabled. Where the substitutions are conservative substitutions, the variants will share immunogenic or antigenic similarity or cross-reactivity with a corresponding natural sequence protein. Natural variants include individual, allelic, polymorphic, strain, or species variants.

Amino acid sequence similarity, or sequence identity, is determined by optimizing residue matches, if necessary, by introducing gaps as required. This changes when considering conservative substitutions as matches. Conservative substitutions typically include substitutions within the following groups: glycine, alanine; valine, isoleucine, leucine; aspartic acid, glutamic acid; asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine. Homologous amino acid sequences include natural allelic and interspecies variations in each respective protein sequence. Typical homologous proteins or peptides will have from 50-100% similarity (if gaps can be introduced), to 75-100% similarity (if conservative substitutions are included) with the amino acid sequence of the relevant DC protein. Identity measures will be at least about 50%, generally at least 60%, more generally at least 65%, usually at least 70%, more usually at least 75%, preferably at least 80%, and more preferably at least

80%, and in particularly preferred embodiments, at least 85% or more. See also Needleham, et al. (1970) J. Mol. Biol. 48:443-453; Sankoff, et al. (1983) Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison Chapter One, Addison-Wesley, Reading, MA; and software packages from the NCBI, at the NIH; and the University of Wisconsin Genetics Computer Group (GCG), Madison, WI.

Nucleic acids encoding the corresponding mammalian DC proteins will typically hybridize to coding portions of SEQ ID NO: 1, 3, 5, or 7 under stringent conditions. For example, nucleic acids encoding the respective DC proteins will typically hybridize to the nucleic acid of SEQ ID NO: 1, 3, 5, or 7, under stringent hybridization conditions, e.g., providing a signal at least 2X background, preferably 5X, 15X, or 25X, while providing few false positive hybridization signals. Generally, stringent conditions are selected to be about 10° C lower than the thermal melting point (T_m) for the sequence being hybridized to at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Typically, stringent conditions will be those in which the salt concentration in wash is about 0.02 molar at pH 7 and the temperature is at least about 50° C. Other factors may significantly affect the stringency of hybridization, including, among others, base composition and size of the complementary strands, the presence of organic solvents such as formamide, and the extent of base mismatching. A preferred embodiment will include nucleic acids which will bind to disclosed sequences in 50% formamide and 20-50 mM NaCl at 42° C.

An isolated DC gene DNA can be readily modified by nucleotide substitutions, nucleotide deletions, nucleotide insertions, and inversions of nucleotide stretches. These modifications result in novel DNA sequences which encode these DC antigens, their derivatives, or proteins having highly similar physiological, immunogenic, or antigenic activity.

Modified sequences can be used to produce mutant antigens or to enhance expression. Enhanced expression may involve gene amplification, increased transcription, increased translation, and other mechanisms. Such mutant DC protein derivatives include predetermined or site-specific mutations of the respective protein or its fragments. "Mutant DC protein" encompasses a polypeptide otherwise falling within the homology definition of the DC protein as set forth above, but having an amino acid sequence which differs from that of the DC protein as found in nature, whether by way of deletion, substitution, or insertion. In particular, "site specific mutant DC protein" generally includes proteins having significant similarity with a protein having a sequence, e.g., of SEQ ID NO: 2. Generally, the variant will share many physicochemical and biological activities, e.g., antigenic or immunogenic, with those sequences, and in preferred embodiments contain most or all of the disclosed sequence. Similar concepts apply to these various DC proteins, particularly those found in various warm blooded animals, e.g., primates and mammals.

Although site specific mutation sites are predetermined, mutants need not be site specific. DC protein mutagenesis can be conducted by making amino acid insertions or deletions. Substitutions, deletions, insertions, or any combinations may be generated to arrive at a final construct. Insertions include amino- or carboxyl- terminal fusions. Random mutagenesis can be conducted at a target codon and the expressed mutants can then be screened for the desired activity. Methods for making substitution mutations at predetermined sites in DNA having a known sequence are well known in the art, e.g., by M13 primer mutagenesis or polymerase chain reaction (PCR) techniques. See also, Sambrook, et al. (1989) and Ausubel, et al. (1987 and Supplements). The mutations in the DNA normally should not place coding sequences out of reading frames and preferably will not create complementary regions that could hybridize to produce secondary mRNA structure such as loops or hairpins.

The present invention also provides recombinant proteins, e.g., heterologous fusion proteins using segments from these proteins. A heterologous fusion protein is a fusion of proteins or segments which are naturally not normally fused in the same manner. Thus, the fusion product of an immunoglobulin with a respective DC polypeptide is a continuous protein molecule having sequences fused in a typical peptide linkage, typically made as a single translation product and exhibiting properties derived from each source peptide. A similar concept applies to heterologous nucleic acid sequences.

In addition, new constructs may be made from combining similar functional domains from other proteins. For example, domains or other segments may be "swapped" between different new fusion polypeptides or fragments, typically with related proteins, e.g., with the lectin or asialoglycoprotein families. Preferably, intact structural domains will be used, e.g., intact Ig portions. See, e.g., Cunningham, et al. (1989) Science 243:1330-1336; and O'Dowd, et al. (1988) J. Biol. Chem. 263:15985-15992. Thus, new chimeric polypeptides exhibiting new combinations of specificities will result from the functional linkage of protein-binding specificities and other functional domains. Also, alanine scanning mutagenesis may be applied, preferably to residues which structurally are exterior to the secondary structure, which will avoid most of the critical residues which generally disrupt tertiary structure.

"Derivatives" of these DC antigens include amino acid sequence mutants, glycosylation variants, and covalent or aggregate conjugates with other chemical moieties. Covalent derivatives can be prepared by linkage of functionalities to groups which are found in these DC protein amino acid side chains or at the N- or C- termini, by means which are well known in the art. These derivatives can include, without limitation, aliphatic esters or amides of the carboxyl terminus, or of residues containing carboxyl side chains, O-acyl derivatives of hydroxyl group-containing residues, and N-

acyl derivatives of the amino terminal amino acid or amino-group containing residues, e.g., lysine or arginine. Acyl groups are selected from the group of alkyl-moieties including C3 to C18 normal alkyl, thereby forming alkanoyl aroyl species. Covalent attachment to carrier proteins may be important when immunogenic moieties are haptens.

In particular, glycosylation alterations are included, e.g., made by modifying the glycosylation patterns of a polypeptide during its synthesis and processing, or in further processing steps. Particularly preferred means for accomplishing this are by exposing the polypeptide to glycosylating enzymes derived from cells which normally provide such processing, e.g., mammalian glycosylation enzymes. Deglycosylation enzymes are also contemplated. Also embraced are versions of the same primary amino acid sequence which have other minor modifications, including phosphorylated amino acid residues, e.g., phosphotyrosine, phosphoserine, or phosphothreonine, or other moieties, including ribosyl groups or cross-linking reagents. Also, proteins comprising substitutions are encompassed, which should retain substantial immunogenicity, to produce antibodies which recognize a protein, e.g., of SEQ ID NO: 2. Typically, these proteins will contain less than 20 residue substitutions from the disclosed sequence, more typically less than 10 substitutions, preferably less than 5, and more preferably less than three. Alternatively, proteins which begin and end at structural domains will usually retain antigenicity and cross immunogenicity.

A major group of derivatives are covalent conjugates of the DC proteins or fragments thereof with other proteins or polypeptides. These derivatives can be synthesized in recombinant culture such as N- or C-terminal fusions or by the use of agents known in the art for their usefulness in cross-linking proteins through reactive side groups. Preferred protein derivatization sites with cross-linking agents are at free amino groups, carbohydrate moieties, and cysteine residues.

Fusion polypeptides between these DC proteins and other homologous or heterologous proteins are also provided. Heterologous polypeptides may be fusions between different surface markers, resulting in, e.g., a hybrid protein. Likewise, heterologous fusions may be constructed which would exhibit a combination of properties or activities of the derivative proteins. Typical examples are fusions of a reporter polypeptide, e.g., luciferase, with a segment or domain of a protein, e.g., a receptor-binding segment, so that the presence or location of the fused protein may be easily determined. See, e.g., Dull, et al., U.S. Patent No. 4,859,609. Other gene fusion partners include bacterial β -galactosidase, trpE, Protein A, β -lactamase, alpha amylase, alcohol dehydrogenase, and yeast alpha mating factor. See, e.g., Godowski, et al. (1988) Science 241:812-816.

Such polypeptides may also have amino acid residues which have been chemically modified by phosphorylation, sulfonation, biotinylation, or the addition or removal of other moieties, particularly those which have molecular shapes similar to phosphate groups. In some embodiments, the modifications will be useful labeling reagents, or serve as purification targets, e.g., affinity ligands.

This invention also contemplates the use of derivatives of these DC proteins other than variations in amino acid sequence or glycosylation. Such derivatives may involve covalent or aggregative association with chemical moieties. These derivatives generally fall into the three classes: (1) salts, (2) side chain and terminal residue covalent modifications, and (3) adsorption complexes, for example with cell membranes. Such covalent or aggregative derivatives are useful as immunogens, as reagents in immunoassays, or in purification methods such as for affinity purification of ligands or other binding ligands. For example, a DC protein antigen can be immobilized by covalent bonding to a solid support such as cyanogen bromide-activated Sepharose, by methods which are well known in the art, or adsorbed onto polyolefin surfaces, with or without

glutaraldehyde cross-linking, for use in the assay or purification of anti-DC protein antibodies. The DC proteins can also be labeled with a detectable group, e.g., radioiodinated by the chloramine T procedure, covalently bound to rare earth chelates, or conjugated to another fluorescent moiety for use in diagnostic assays. Purification of these SDCMP proteins may be effected by immobilized antibodies.

Isolated DC protein genes will allow transformation of cells lacking expression of a corresponding DC protein, e.g., either species types or cells which lack corresponding proteins and exhibit negative background activity. Expression of transformed genes will allow isolation of antigenically pure cell lines, with defined or single specie variants. This approach will allow for more sensitive detection and discrimination of the physiological effects of these DC proteins. Subcellular fragments, e.g., cytoplasts or membrane fragments, can be isolated and used.

VIII. Binding Agent:DC Protein Complexes

A DC protein that specifically binds to or that is specifically immunoreactive with an antibody generated against a defined immunogen, e.g., an immunogen consisting of the amino acid sequence of SEQ ID NO: 2, is determined in an immunoassay. The immunoassay uses a polyclonal antiserum which was raised to the protein of SEQ ID NO: 2. This antiserum is selected to have low crossreactivity against other members of the related families, and any such crossreactivity is removed by immunoabsorption prior to use in the immunoassay.

In order to produce antisera for use in an immunoassay, e.g., the protein of SEQ ID NO: 2, is isolated as described herein. For example, recombinant protein may be produced in a mammalian cell line. An inbred strain of mice such as BALB/c is immunized with the appropriate protein using a standard adjuvant, such as Freund's adjuvant, and a standard mouse immunization protocol (see Harlow and Lane, supra). Alternatively, a synthetic peptide derived from the sequences disclosed

herein and conjugated to a carrier protein can be used an immunogen. Polyclonal sera are collected and titrated against the immunogen protein in an immunoassay, e.g., a solid phase immunoassay with the immunogen immobilized on a solid support. Polyclonal antisera with a titer of 10^4 or greater are selected and tested for their cross reactivity against other related proteins, using a competitive binding immunoassay such as the one described in Harlow and Lane, supra, at pages 570-573. Preferably two different related proteins are used in this determination in conjunction with a given DC protein. For example, with the lectin protein, at least two other family members are used to absorb out shared epitopes. In conjunction with the SDCMP3 family member, two other members of the family are used. These other family members can be produced as recombinant proteins and isolated using standard molecular biology and protein chemistry techniques as described herein.

Immunoassays in the competitive binding format can be used for the crossreactivity determinations. For example, the protein of SEQ ID NO: 2 can be immobilized to a solid support. Proteins added to the assay compete with the binding of the antisera to the immobilized antigen. The ability of the above proteins to compete with the binding of the antisera to the immobilized protein is compared to the protein of SEQ ID NO 2. The percent crossreactivity for the above proteins is calculated, using standard calculations. Those antisera with less than 10% crossreactivity with each of the proteins listed above are selected and pooled. The cross-reacting antibodies are then removed from the pooled antisera by immunoabsorption with the above-listed proteins.

The immunoabsorbed and pooled antisera are then used in a competitive binding immunoassay as described above to compare a second protein to the immunogen protein (e.g., the SDCMP3 protein of SEQ ID NO: 2). In order to make this comparison, the two proteins are each assayed at a wide range of concentrations and the amount of each protein required to inhibit 50% of the binding of the

antisera to the immobilized protein is determined. If the amount of the second protein required is less than twice the amount of the protein of SEQ ID NO: 2 that is required, then the second protein is said to specifically bind to an antibody generated to the immunogen.

It is understood that DC proteins are likely a family of homologous proteins that comprise two or more genes. For a particular gene product, such as the human Ig family member protein, the invention encompasses not only the amino acid sequences disclosed herein, but also to other proteins that are allelic, polymorphic, non-allelic, or species variants. It is also understood that the term "human DC protein" includes nonnatural mutations introduced by deliberate mutation using conventional recombinant technology such as single site mutation, or by excising short sections of DNA encoding these proteins or splice variants from the gene, or by substituting or adding small numbers of new amino acids. Such minor alterations must substantially maintain the immunoidentity of the original molecule and/or its biological activity. Thus, these alterations include proteins that are specifically immunoreactive with a designated naturally occurring respective SDCMP protein, e.g., the human SDCMP4 protein exhibiting SEQ ID NO: 6 or 8. Particular protein modifications considered minor would include conservative substitution of amino acids with similar chemical properties, as described above for each protein family as a whole. By aligning a protein optimally with the protein of SEQ ID NO 2, and by using the conventional immunoassays described herein to determine immunoidentity, one can determine the protein compositions of the invention.

IX. Uses

The present invention provides reagents which will find use in diagnostic applications as described elsewhere herein, e.g., in the general description for developmental abnormalities, or below in the description of kits for diagnosis.

DC genes, e.g., DNA or RNA may be used as a component in a forensic assay. For instance, the nucleotide sequences provided may be labeled using, e.g., ³²P or biotin and used to probe standard restriction fragment polymorphism blots, providing a measurable character to aid in distinguishing between individuals. Such probes may be used in well-known forensic techniques such as genetic fingerprinting. In addition, nucleotide probes made from DC sequences may be used in in situ assays to detect chromosomal abnormalities.

Antibodies and other binding agents directed towards DC proteins or nucleic acids may be used to purify the corresponding DC protein molecule. As described in the Examples below, antibody purification of DC proteins is both possible and practicable. Antibodies and other binding agents may also be used in a diagnostic fashion to determine whether DC components are present in a tissue sample or cell population using well-known techniques described herein. The ability to attach a binding agent to a DC protein provides a means to diagnose disorders associated with expression misregulation. Antibodies and other DC protein binding agents may also be useful as histological markers, or purification reagents. As described in the examples below, the expression of each of these proteins is limited to specific tissue types. By directing a probe, such as an antibody or nucleic acid to the respective DC protein, it is possible to use the probe to distinguish tissue and cell types in situ or in vitro.

In addition, purified antigen may be used to deplete an antiserum preparation of those antibodies which bind with selectivity to the antigen. Thus, e.g., the mouse SDCMP3 may be used to deplete an antiserum raised to human SDCMP4 of components which may cross react with mouse SDCMP3. Alternatively, the SDCMP3 may be used to purify those components of an antiserum which bind with affinity to the respective antigen.

SDCMP4 shares a number of features with the hepatic ASGPR, the best known example of the type II transmembrane C-type lectins. The hepatic ASGPR displays

binding specificity for galactose residues, and its intracellular domain bears a tyrosine motif for ligand internalization. These features enable the hepatic ASGPR to bind desialylated plasma glycoproteins expressing galactose residues, and subsequently provide for clearance of those proteins from the plasma.

The ligand specificity of SDCMP4 cannot be absolutely inferred from its CRD sequence. However, the expression of SDCMP4 on DC is an indication that potentially antigenic constituents, such as found on microorganisms, could represent natural ligands of SDCMP4. In this context, the mannose-receptor, another C-type lectin found on DC and macrophages, has the capacity to bind and internalize, e.g., yeast particles following recognition of the mannose moieties of their cell wall.

The presence of a tyrosine-based internalization motif in SDCMP4 predicts that the molecule plays a role in receptor-mediated endocytosis by DC. It can be suggested that SDCMP4 functions as an "antigen-receptor" in DC, to internalize ligands that will subsequently be routed into an intracellular processing pathway resulting in antigen presentation and initiation or promotion of an immune response.

Such an internalization function mediated by SDCMP4 makes this receptor a potential target for directing antigens into DC, e.g., for enhancing presentation to T cells, and subsequent activation of specific immunity. Thus, SDCMP4 could represent a receptor for delivery of antigen in vaccination protocols, thereby targeting the antigen to the appropriate cells for initiation of a vaccine response. The therapeutic significance of such strategy might be of particular relevance in cancer immunotherapy, where tumor-associated antigens (TAA) could be coupled to reagents specifically recognizing SDCMP4 for selective delivery to DC.

This invention also provides reagents which may exhibit significant therapeutic value. The DC proteins (naturally occurring or recombinant), fragments thereof, and antibodies thereto, along with compounds identified

as having binding affinity to the DC protein, may be useful in the treatment of conditions associated with abnormal physiology or development, including abnormal proliferation, e.g., cancerous conditions, or
5 degenerative conditions. Abnormal proliferation, regeneration, degeneration, and atrophy may be modulated by appropriate therapeutic treatment using the compositions provided herein. For example, a disease or disorder associated with abnormal expression or abnormal
10 signaling by a DC, e.g., as an antigen presenting cell, is a target for an agonist or antagonist of the protein. The proteins likely play a role in regulation or development of hematopoietic cells, e.g., lymphoid cells, which affect immunological responses, e.g., antigen
15 presentation and the resulting effector functions.

It is believed that blocking the interaction of these SDCMPs may block signaling. Thus, e.g., use of polyclonal or selected monoclonal antibodies against the proteins may affect immune responses, e.g., MLR.
20 Alternatively, soluble extracellular fragments may block interaction with a counterreceptor, thus also blocking such a reaction. Since MLR is diagnostic of initiation or maintenance of an immune response, these reagents may be useful in modulating the initiation and maintenance of
25 immune responses.

Other abnormal developmental conditions are known in cell types shown to possess DC protein mRNA by northern blot analysis. See Berkow (ed.) The Merck Manual of Diagnosis and Therapy, Merck and Co., Rahway, NJ; and
30 Thorn, et al. Harrison's Principles of Internal Medicine, McGraw-Hill, NY. Developmental or functional abnormalities, e.g., of the immune system, cause significant medical abnormalities and conditions which may be susceptible to prevention or treatment using
35 compositions provided herein.

Recombinant DC proteins or antibodies might be purified and then administered to a patient. These reagents can be combined for therapeutic use with additional active or inert ingredients, e.g., in
40 conventional pharmaceutically acceptable carriers or

diluents, e.g., immunogenic adjuvants, along with physiologically innocuous stabilizers and excipients. In particular, these may be useful in a vaccine context, where the antigen is combined with one of these therapeutic versions of agonists or antagonists. These combinations can be sterile filtered and placed into dosage forms as by lyophilization in dosage vials or storage in stabilized aqueous preparations. This invention also contemplates use of antibodies or binding fragments thereof, including forms which are not complement binding.

Drug screening using antibodies or receptor or fragments thereof can identify compounds having binding affinity to these DC proteins, including isolation of associated components. Subsequent biological assays can then be utilized to determine if the compound has intrinsic stimulating activity and is therefore a blocker or antagonist in that it blocks the activity of the protein. Likewise, a compound having intrinsic stimulating activity might activate the cell through the protein and is thus an agonist in that it simulates the cell. This invention further contemplates the therapeutic use of antibodies to the proteins as antagonists.

The quantities of reagents necessary for effective therapy will depend upon many different factors, including means of administration, target site, physiological state of the patient, and other medicants administered. Thus, treatment dosages should be titrated to optimize safety and efficacy. Typically, dosages used in vitro may provide useful guidance in the amounts useful for in situ administration of these reagents. Animal testing of effective doses for treatment of particular disorders will provide further predictive indication of human dosage. Various considerations are described, e.g., in Gilman, et al. (eds.) (1990) Goodman and Gilman's: The Pharmacological Bases of Therapeutics (8th ed.) Pergamon Press; and (1990) Remington's Pharmaceutical Sciences (17th ed.) Mack Publishing Co., Easton, PA. Methods for administration are discussed

therein and below, e.g., for oral, intravenous, intraperitoneal, or intramuscular administration, transdermal diffusion, and others. Pharmaceutically acceptable carriers will include water, saline, buffers, and other compounds described, e.g., in the Merck Index, Merck and Co., Rahway, NJ. Dosage ranges would ordinarily be expected to be in amounts lower than 1 mM concentrations, typically less than about 10 μ M concentrations, usually less than about 100 nM, preferably less than about 10 pM (picomolar), and most preferably less than about 1 fM (femtomolar), with an appropriate carrier. Slow release formulations, or a slow release apparatus will often be utilized for continuous administration.

The DC proteins, fragments thereof, and antibodies to it or its fragments, antagonists, and agonists, could be administered directly to the host to be treated or, depending on the size of the compounds, it may be desirable to conjugate them to carrier proteins such as ovalbumin or serum albumin prior to their administration. Therapeutic formulations may be administered in many conventional dosage formulations. While it is possible for the active ingredient to be administered alone, it is preferable to present it as a pharmaceutical formulation. Formulations typically comprise at least one active ingredient, as defined above, together with one or more acceptable carriers thereof. Each carrier should be both pharmaceutically and physiologically acceptable in the sense of being compatible with the other ingredients and not injurious to the patient. Formulations include those suitable for oral, rectal, nasal, or parenteral (including subcutaneous, intramuscular, intravenous and intradermal) administration. The formulations may conveniently be presented in unit dosage form and may be prepared by any methods well known in the art of pharmacy. See, e.g., Gilman, et al. (eds.) (1990) Goodman and Gilman's: The Pharmacological Bases of Therapeutics (8th ed.) Pergamon Press; and (1990) Remington's Pharmaceutical Sciences (17th ed.) Mack Publishing Co., Easton, PA; Avis, et al. (eds.) (1993)

Pharmaceutical Dosage Forms: Parenteral Medications

Dekker, NY; Lieberman, et al. (eds.) (1990)

Pharmaceutical Dosage Forms: Tablets Dekker, NY; and

Lieberman, et al. (eds.) (1990) Pharmaceutical Dosage

5 Forms: Disperse Systems Dekker, NY. The therapy of this invention may be combined with or used in association with other chemotherapeutic or chemopreventive agents.

Both the naturally occurring and the recombinant form of the DC proteins of this invention are
10 particularly useful in kits and assay methods which are capable of screening compounds for binding activity to the proteins. Several methods of automating assays have been developed in recent years so as to permit screening of tens of thousands of compounds in a short period.
15 See, e.g., Fodor, et al. (1991) Science 251:767-773, and other descriptions of chemical diversity libraries, which describe means for testing of binding affinity by a plurality of compounds. The development of suitable assays can be greatly facilitated by the availability of
20 large amounts of purified, e.g., soluble versions of, DC protein as provided by this invention.

For example, antagonists can often be found once the protein has been structurally defined. Testing of potential protein analogs is now possible upon the
25 development of highly automated assay methods using a purified surface protein. In particular, new agonists and antagonists will be discovered by using screening techniques described herein. Of particular importance are compounds found to have a combined binding affinity
30 for multiple related cell surface antigens, e.g., compounds which can serve as antagonists for species variants of a DC protein.

This invention is particularly useful for screening compounds by using recombinant DC protein in a variety of
35 drug screening techniques. The advantages of using a recombinant protein in screening for specific ligands include: (a) improved renewable source of the protein from a specific source; (b) potentially greater number of antigens per cell giving better signal to noise ratio in

assays; and (c) species variant specificity (theoretically giving greater biological and disease specificity).

5 One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant DNA molecules expressing a DC protein. Cells may be isolated which express that protein in isolation from any others. Such cells, either in viable or fixed form, can be used for standard surface protein binding assays. See also, Parce, et al. (1989) Science 246:243-10 247; and Owicki, et al. (1990) Proc. Nat'l Acad. Sci. USA 87:4007-4011, which describe sensitive methods to detect cellular responses. Competitive assays are particularly useful, where the cells (source of DC protein) are 15 contacted and incubated with an antibody having known binding affinity to the antigen, such as ^{125}I -antibody, and a test sample whose binding affinity to the binding composition is being measured. The bound and free labeled binding compositions are then separated to assess the degree of protein binding. The amount of test 20 compound bound is inversely proportional to the amount of labeled antibody binding to the known source. Many techniques can be used to separate bound from free reagent to assess the degree of binding. This separation step could typically involve a procedure such as adhesion 25 to filters followed by washing, adhesion to plastic followed by washing, or centrifugation of the cell membranes. Viable cells could also be used to screen for the effects of drugs on these DC protein mediated functions, e.g., antigen presentation or helper function. 30

Another method utilizes membranes from transformed eukaryotic or prokaryotic host cells as the source of a DC protein. These cells are stably transformed with DNA vectors directing the expression of the appropriate 35 protein, e.g., an engineered membrane bound form. Essentially, the membranes would be prepared from the cells and used in binding assays such as the competitive assay set forth above.

40 Still another approach is to use solubilized, unpurified or solubilized, purified DC protein from

transformed eukaryotic or prokaryotic host cells. This allows for a "molecular" binding assay with the advantages of increased specificity, the ability to automate, and high drug test throughput.

5 Another technique for drug screening involves an approach which provides high throughput screening for compounds having suitable binding affinity to the respective DC protein and is described in detail in Geysen, European Patent Application 84/03564, published
10 on September 13, 1984. First, large numbers of different small peptide test compounds are synthesized on a solid substrate, e.g., plastic pins or some other appropriate surface, see Fodor, et al., supra. Then all the pins are reacted with solubilized, unpurified or solubilized,
15 purified DC protein, and washed. The next step involves detecting bound reagent, e.g., antibody.

 One means for determining which sites interact with specific other proteins is a physical structure determination, e.g., x-ray crystallography or 2
20 dimensional NMR techniques. These will provide guidance as to which amino acid residues form molecular contact regions. For a detailed description of protein structural determination, see, e.g., Blundell and Johnson (1976) Protein Crystallography Academic Press, NY.

25

X. Kits

 This invention also contemplates use of these DC proteins, fragments thereof, peptides, and their fusion products in a variety of diagnostic kits and methods for
30 detecting the presence of a DC protein or message. Typically the kit will have a compartment containing either a defined DC peptide or gene segment or a reagent which recognizes one or the other, e.g., antibodies.

 A kit for determining the binding affinity of a test
35 compound to the respective DC protein would typically comprise a test compound; a labeled compound, for example an antibody having known binding affinity for the protein; a source of the DC protein (naturally occurring or recombinant); and a means for separating bound from
40 free labeled compound, such as a solid phase for

immobilizing the DC protein. Once compounds are screened, those having suitable binding affinity to the protein can be evaluated in suitable biological assays, as are well known in the art, to determine whether they act as agonists or antagonists to regulate DC function. The availability of recombinant DC polypeptides also provide well defined standards for calibrating such assays.

A preferred kit for determining the concentration of, for example, a DC protein in a sample would typically comprise a labeled compound, e.g., antibody, having known binding affinity for the DC protein, a source of DC protein (naturally occurring or recombinant) and a means for separating the bound from free labeled compound, for example, a solid phase for immobilizing the DC protein. Compartments containing reagents, and instructions, will normally be provided.

Antibodies, including antigen binding fragments, specific for the respective DC or its fragments are useful in diagnostic applications to detect the presence of elevated levels of the protein and/or its fragments. Such diagnostic assays can employ lysates, live cells, fixed cells, immunofluorescence, cell cultures, body fluids, and further can involve the detection of antigens in serum, or the like. Diagnostic assays may be homogeneous (without a separation step between free reagent and antigen-DC protein complex) or heterogeneous (with a separation step). Various commercial assays exist, such as radioimmunoassay (RIA), enzyme-linked immunosorbent assay (ELISA), enzyme immunoassay (EIA), enzyme-multiplied immunoassay technique (EMIT), substrate-labeled fluorescent immunoassay (SLFIA), and the like. For example, unlabeled antibodies can be employed by using a second antibody which is labeled and which recognizes the antibody to the DC protein or to a particular fragment thereof. Similar assays have also been extensively discussed in the literature. See, e.g., Harlow and Lane (1988) Antibodies: A Laboratory Manual, CSH Press, NY; Chan (ed. 1987) Immunoassay: A Practical Guide Academic Press, Orlando, FL; Price and Newman (eds.

1991) Principles and Practice of Immunoassay Stockton Press, NY; and Ngo (ed. 1988) Nonisotopic Immunoassay Plenum Press, NY. In particular, the reagents may be useful for diagnosing DC populations in biological samples, either to detect an excess or deficiency of DC in a sample. The assay may be directed to histological analysis of a biopsy, or evaluation of DC numbers in a blood or tissue sample.

Anti-idiotypic antibodies may have similar use to diagnose presence of antibodies against a DC protein, as such may be diagnostic of various abnormal states. For example, overproduction of the DC protein may result in various immunological reactions which may be diagnostic of abnormal physiological states, particularly in proliferative cell conditions such as cancer or abnormal differentiation.

Frequently, the reagents for diagnostic assays are supplied in kits, so as to optimize the sensitivity of the assay. For the subject invention, depending upon the nature of the assay, the protocol, and the label, either labeled or unlabeled antibody or receptor, or labeled DC protein is provided. This is usually in conjunction with other additives, such as buffers, stabilizers, materials necessary for signal production such as substrates for enzymes, and the like. Preferably, the kit will also contain instructions for proper use and disposal of the contents after use. Typically the kit has compartments for each useful reagent. Desirably, the reagents are provided as a dry lyophilized powder, where the reagents may be reconstituted in an aqueous medium providing appropriate concentrations of reagents for performing the assay.

Many of the aforementioned constituents of the drug screening and the diagnostic assays may be used without modification or may be modified in a variety of ways. For example, labeling may be achieved by covalently or non-covalently joining a moiety which directly or indirectly provides a detectable signal. In many of these assays, the protein, test compound, DC protein, or

antibodies thereto can be labeled either directly or indirectly. Possibilities for direct labeling include label groups: radiolabels such as ^{125}I , enzymes (U.S. Pat. No. 3,645,090) such as peroxidase and alkaline phosphatase, and fluorescent labels (U.S. Pat. No. 3,940,475) capable of monitoring the change in fluorescence intensity, wavelength shift, or fluorescence polarization. Possibilities for indirect labeling include biotinylation of one constituent followed by binding to avidin coupled to one of the above label groups.

There are also numerous methods of separating the bound from the free protein, or alternatively the bound from the free test compound. The DC protein can be immobilized on various matrices followed by washing. Suitable matrices include plastic such as an ELISA plate, filters, and beads. Methods of immobilizing the DC protein to a matrix include, without limitation, direct adhesion to plastic, use of a capture antibody, chemical coupling, and biotin-avidin. The last step in this approach involves the precipitation of protein/antibody complex by one of several methods including those utilizing, e.g., an organic solvent such as polyethylene glycol or a salt such as ammonium sulfate. Other suitable separation techniques include, without limitation, the fluorescein antibody magnetizable particle method described in Rattle, et al. (1984) Clin. Chem. 30:1457-1461, and the double antibody magnetic particle separation as described in U.S. Pat. No. 4,659,678.

Methods for linking proteins or their fragments to the various labels have been extensively reported in the literature and do not require detailed discussion here. Many of the techniques involve the use of activated carboxyl groups either through the use of carbodiimide or active esters to form peptide bonds, the formation of thioethers by reaction of a mercapto group with an activated halogen such as chloroacetyl, or an activated olefin such as maleimide, for linkage, or the like. Fusion proteins will also find use in these applications.

Another diagnostic aspect of this invention involves use of oligonucleotide or polynucleotide sequences taken from the sequence of a respective DC protein. These sequences can be used as probes for detecting levels of the message in samples from patients suspected of having an abnormal condition, e.g., cancer or immune problem. The preparation of both RNA and DNA nucleotide sequences, the labeling of the sequences, and the preferred size of the sequences has received ample description and discussion in the literature. Normally an oligonucleotide probe should have at least about 14 nucleotides, usually at least about 18 nucleotides, and the polynucleotide probes may be up to several kilobases. Various labels may be employed, most commonly radionuclides, particularly ^{32}P . However, other techniques may also be employed, such as using biotin modified nucleotides for introduction into a polynucleotide. The biotin then serves as the site for binding to avidin or antibodies, which may be labeled with a wide variety of labels, such as radionuclides, fluorophores, enzymes, or the like. Alternatively, antibodies may be employed which can recognize specific duplexes, including DNA duplexes, RNA duplexes, DNA-RNA hybrid duplexes, or DNA-protein duplexes. The antibodies in turn may be labeled and the assay carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected. The use of probes to the novel anti-sense RNA may be carried out in any conventional techniques such as nucleic acid hybridization, plus and minus screening, recombinational probing, hybrid released translation (HRT), and hybrid arrested translation (HART). This also includes amplification techniques such as polymerase chain reaction (PCR).

Diagnostic kits which also test for the qualitative or quantitative presence of other markers are also contemplated. Diagnosis or prognosis may depend on the combination of multiple indications used as markers. Thus, kits may test for combinations of markers. See,

e.g., Viallet, et al. (1989) Progress in Growth Factor Res. 1:89-97.

XI. Binding Partner Isolation

5 Having isolated one member of a binding partner of a specific interaction, methods exist for isolating the counter-partner. See, Gearing, et al. (1989) EMBO J. 8:3667-3676. For example, means to label a DC surface protein without interfering with the binding to its
10 receptor can be determined. For example, an affinity label can be fused to either the amino- or carboxyl-terminus of the ligand. An expression library can be screened for specific binding to the DC protein, e.g., by cell sorting, or other screening to detect subpopulations
15 which express such a binding component. See, e.g., Ho, et al. (1993) Proc. Nat'l Acad. Sci. USA 90:11267-11271. Alternatively, a panning method may be used. See, e.g., Seed and Aruffo (1987) Proc. Nat'l Acad. Sci. USA 84:3365-3369. A two-hybrid selection system may also be
20 applied making appropriate constructs with the available DC protein sequences. See, e.g., Fields and Song (1989) Nature 340:245-246.

 Protein cross-linking techniques with label can be applied to isolate binding partners of a DC protein.
25 This would allow identification of proteins which specifically interact with the appropriate DC protein.

 The broad scope of this invention is best understood with reference to the following examples, which are not intended to limit the invention to specific embodiments.

30

EXAMPLES

I. General Methods

 Many of the standard methods below are described or referenced, e.g., in Maniatis, et al. (1982) Molecular Cloning, A Laboratory Manual Cold Spring Harbor
35 Laboratory, Cold Spring Harbor Press, NY; Sambrook, et al. (1989) Molecular Cloning: A Laboratory Manual (2d ed.) Vols. 1-3, CSH Press, NY; Ausubel, et al., Biology Greene Publishing Associates, Brooklyn, NY; or Ausubel,

et al. (1987 and Supplements) Current Protocols in Molecular Biology Wiley/Greene, NY; Innis, et al. (eds.) (1990) PCR Protocols: A Guide to Methods and Applications Academic Press, NY.

5 Methods for protein purification include such methods as ammonium sulfate precipitation, column chromatography, electrophoresis, centrifugation, crystallization, and others. See, e.g., Ausubel, et al. (1987 and periodic supplements); Deutscher (1990) "Guide to Protein Purification," Methods in Enzymology vol. 182, 10 and other volumes in this series; Coligan, et al. (1996 and periodic Supplements) Current Protocols in Protein Science Wiley/Greene, NY; and manufacturer's literature on use of protein purification products, e.g., Pharmacia, 15 Piscataway, NJ, or Bio-Rad, Richmond, CA. Combination with recombinant techniques allow fusion to appropriate segments, e.g., to a FLAG sequence or an equivalent which can be fused via a protease-removable sequence. See, e.g., Hochuli (1989) Chemische Industrie 12:69-70; 20 Hochuli (1990) "Purification of Recombinant Proteins with Metal Chelate Absorbent" in Setlow (ed.) Genetic Engineering, Principle and Methods 12:87-98, Plenum Press, NY; and Crowe, et al. (1992) OIAexpress: The High Level Expression and Protein Purification System QUIAGEN, 25 Inc., Chatsworth, CA.

 Methods for determining immunological function are described, e.g., in Hertenberg, et al. (eds. 1996) Weir's Handbook of Experimental Immunology vols. 1-4, Blackwell Science; Coligan, et al. (1992 and periodic 30 Supplements) Current Protocols in Immunology Wiley/Greene, NY; and Methods in Enzymology volumes. 70, 73, 74, 84, 92, 93, 108, 116, 121, 132, 150, 162, and 163. See also, e.g., Paul (ed.) (1993) Fundamental Immunology (3d ed.) Raven Press, N.Y. Particularly 35 useful functions of dendritic cells are described, e.g., in Steinman (1991) Annual Review of Immunology 9:271-296; and Banchereau and Schmitt (eds. 1994) Dendritic Cells in Fundamental and Clinical Immunology Plenum Press, NY.

FACS analyses are described in Melamed, et al. (1990) Flow Cytometry and Sorting Wiley-Liss, Inc., New York, NY; Shapiro (1988) Practical Flow Cytometry Liss, New York, NY; and Robinson, et al. (1993) Handbook of Flow Cytometry Methods Wiley-Liss, New York, NY.

II. Generation of dendritic cells

Human CD34+ cells were obtained as follows. See, e.g., Caux, et al. (1995) pages 1-5 in Banchereau and Schmitt Dendritic Cells in Fundamental and Clinical Immunology Plenum Press, NY. Peripheral or cord blood cells, sometimes CD34+ selected, were cultured in the presence of Stem Cell Factor (SCF), GM-CSF, and TNF- α in endotoxin free RPMI 1640 medium (GIBCO, Grand Island, NY) supplemented with 10% (v/v) heat-inactivated fetal bovine serum (FBS; Flow Laboratories, Irvine, CA), 10 mM HEPES, 2 mM L-glutamine, 5×10^{-5} M 2-mercaptoethanol, penicillin (100 μ g/ml). This is referred to as complete medium.

CD34+ cells were seeded for expansion in 25 to 75 cm² flasks (Corning, NY) at 2×10^4 cells/ml. Optimal conditions were maintained by splitting these cultures at day 5 and 10 with medium containing fresh GM-CSF and TNF- α (cell concentration: $1-3 \times 10^5$ cells/ml). In certain cases, cells were FACS sorted for CD1a expression at about day 6.

In certain situations, cells were routinely collected after 12 days of culture, eventually adherent cells were recovered using a 5 mM EDTA solution. In other situations, the CD1a+ cells were activated by resuspension in complete medium at 5×10^6 cells/ml and activated for the appropriate time (e.g., 1 or 6 h) with 1 μ g/ml phorbol 12-myristate 13-acetate (PMA, Sigma) and 100 ng/ml ionomycin (Calbiochem, La Jolla, CA). These cells were expanded for another 6 days, and RNA isolated for cDNA library preparation.

III. RNA isolation and library construction

Total RNA is isolated using, e.g., the guanidine thiocyanate/CsCl gradient procedure as described by Chirgwin, et al. (1978) Biochem. 18:5294-5299.

5 Alternatively, poly(A)+ RNA is isolated using the OLIGOTEX mRNA isolation kit (QIAGEN). Double stranded cDNA are generated using, e.g., the SUPERScript plasmid system (Gibco BRL, Gaithersburg, MD) for cDNA synthesis and plasmid cloning. The resulting double stranded cDNA is unidirectionally cloned, e.g., into pSport1 and
10 transfected by electroporation into ELECTROMAX DH10BTM Cells (Gibco BRL, Gaithersburg, MD).

IV. Sequencing

DNA isolated from randomly picked clones, or after
15 subtractive hybridization using unactivated cells, were subjected to nucleotide sequence analysis using standard techniques. A Taq DiDeoxy Terminator cycle sequencing kit (Applied Biosystems, Foster City, CA) can be used. The labeled DNA fragments are separated using a DNA
20 sequencing gel of an appropriate automated sequencer. Alternatively, the isolated clone is sequenced as described, e.g., in Maniatis, et al. (1982) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor Press; Sambrook, et al.
25 (1989) Molecular Cloning: A Laboratory Manual, (2d ed.), vols. 1-3, CSH Press, NY; Ausubel, et al., Biology, Greene Publishing Associates, Brooklyn, NY; or Ausubel, et al. (1987 and Supplements) Current Protocols in Molecular Biology, Greene/Wiley, New York. Chemical
30 sequencing methods are also available, e.g., using Maxam and Gilbert sequencing techniques.

V. Recombinant DC gene construct

Poly(A)+ RNA is isolated from appropriate cell
35 populations, e.g., using the FastTrack mRNA kit (Invitrogen, San Diego, CA). Samples are electrophoresed, e.g., in a 1% agarose gel containing formaldehyde and transferred to a GeneScreen membrane (NEN Research Products, Boston, MA). Hybridization is
40 performed, e.g., at 65° C in 0.5 M NaHPO₄ pH 7.2, 7% SDS,

1 mM EDTA, and 1% BSA (fraction V) with ^{32}P -dCTP labeled DC gene cDNA at 10^7 cpm/ml. After hybridization, filters are washed three times at 50° C in 0.2X SSC, 0.1% SDS, e.g., for 30 min, and exposed to film for 24 h. A
5 positive signal will typically be 2X over background, preferably 5-25X.

The recombinant gene construct may be used to generate a probe for detecting the message. The insert may be excised and used in the detection methods
10 described above. Various standard methods for cross species hybridization and washes are well known in the art. See, e.g., Sambrook, et al. and Ausubel.

VI. Expression of DC gene Protein in E. coli.

15 PCR is used to make a construct comprising the open reading frame, preferably in operable association with proper promoter, selection, and regulatory sequences. The resulting expression plasmid is transformed into an appropriate, e.g., the Topp5, E. coli strain (Stratagene,
20 La Jolla, CA). Ampicillin resistant (50 µg/ml) transformants are grown in Luria Broth (Gibco) at 37° C until the optical density at 550 nm is 0.7. Recombinant protein is induced with 0.4 mM isopropyl-βD-thiogalactopyranoside (Sigma, St. Louis, MO) and incubation of the
25 cells continued at 20° C for a further 18 hours. Cells from a 1 liter culture are harvested by centrifugation and resuspended, e.g., in 200 ml of ice cold 30% sucrose, 50 mM Tris HCl pH 8.0, 1 mM ethylenediaminetetraacetic acid. After 10 min on ice, ice cold water is added to a
30 total volume of 2 liters. After 20 min on ice, cells are removed by centrifugation and the supernatant is clarified by filtration via a 5 µM Millipak 60 (Millipore Corp., Bedford, MA).

The recombinant protein is purified via standard
35 purification methods, e.g., various ion exchange chromatography methods. Immunoaffinity methods using antibodies described below can also be used. Affinity methods may be used where an epitope tag is engineered into an expression construct.

Similar methods are used to prepare expression constructs and cells in eukaryotic cells. Eukaryotic promoters and expression vectors may be produced, as described above.

5

VII. Mapping of human DC genes

DNA isolation, restriction enzyme digestion, agarose gel electrophoresis, Southern blot transfer and hybridization are performed according to standard techniques. See Jenkins, et al. (1982) J. Virol. 43:26-36. Blots may be prepared with Hybond-N nylon membrane (Amersham). The probe is labeled with ^{32}P -dCTP; washing is done to a final stringency, e.g., of 0.1X SSC, 0.1% SDS, 65° C.

10

Alternatively, a BIOS Laboratories (New Haven, CT) mouse somatic cell hybrid panel may be combined with PCR methods. See Fan, et al. (1996) Immunogenetics 44:97-103.

15

The human SDCMP3 gene is localized at chromosome 12 p12-13 (human NK receptor complex), as determined by radiation hybrid mapping with PCR primers.

VIII. Analysis of individual variation

20

From the distribution data, an abundant easily accessible cell type is selected for sampling from individuals. Using PCR techniques, a large population of individuals are analyzed for this gene. cDNA or other PCR methods are used to sequence the corresponding gene in the different individuals, e.g., outbred mouse strains, and their sequences are compared. This indicates both the extent of divergence among racial or other populations, as well as determining which residues are likely to be modifiable without dramatic effects on function.

25

IX. Preparation of Antibodies

Recombinant DC proteins are generated by expression in *E. coli* as shown above, and tested for biological activity. Alternatively, natural protein sources may be used with purification methods made available. Antibody

30

reagents may be used in immunopurification, or to track separation methods. Active or denatured proteins may be used for immunization of appropriate mammals for either polyclonal serum production, or for monoclonal antibody production.

X. Isolation of counterpart DC genes

Human cDNA clones encoding these genes are used as probes, or to design PCR primers, to find counterparts in various primate species, e.g., chimpanzees. Others may be identified from other animals, e.g., domesticated farm or pet animal species.

XI. Use of reagents to analyze cell populations

Detection of the level of dendritic cells present in a sample is important for diagnosis of aberrant disease conditions. For example, an increase in the number of dendritic cells in a tissue or the lymph system can be indicative of the presence of a DC hyperplasia, or tissue or graft rejection. A low DC population can indicate an abnormal reaction to, e.g., a bacterial or viral infection, which may require the appropriate treat to normalize the DC response.

FACS analysis using a labeled binding agent specific for a cell surface DC protein, see, e.g., Melamed, et al. (1990) Flow Cytometry and Sorting Wiley-Liss, Inc., New York, NY; Shapiro (1988) Practical Flow Cytometry Liss, New York, NY; and Robinson, et al. (1993) Handbook of Flow Cytometry Methods Wiley-Liss, New York, NY, is used in determining the number of DCs present in a cell mixture, e.g., PBMCs, adherent cells, etc. The binding agent is also used for histological analysis of tissue samples, either fresh or fixed, to analyze infiltration of DC. Diverse cell populations may also be evaluated, either in a cell destructive assay, or in certain assays where cells retain viability. Alternatively, tissue or cell fixation methods may be used.

Levels of DC transcripts are quantitated, e.g., using semiquantitative PCR as described in Murphy, et al. (1993) J. Immunol. Methods 162:211-223. Primers or other

methods are designed such that genomic DNA is not detected.

XII. Preparing Immunoselective binding preparations

5 Polyclonal antiserum is prepared, e.g., as described above. The other asialoglycoprotein receptors are used to deplete components which bind specifically to them, leaving components which will bind to the desired SDCMP3 or SDCMP4. Such depleted sera can be linked to a solid
10 substrate, e.g., and used to immunoselect the antigen from an impure source. Immunoselected antigen may be subject to further purification by standard protein purification procedures, e.g., ammonium sulfate precipitations, ion exchange, or other chromatography
15 methods, HPLC, etc. The specific serum may be used to follow the purification, e.g., determining what fractions the desired protein partitions.

XIII. Expression distribution

20 Two transcripts have been detected for the human SDCMP3 by PCR analysis. The shorter form corresponds to a form where a deletion corresponding to nucleotides 376-513 (269-406 of ORF), which retains the open reading frame. The cellular distribution of the two forms appear
25 to be similar.

The distribution of the primate SDCMP3 is detected in DC prepared from CD34+ progenitors cultured 12 d in GM-CSF and TNF α , activated 1-6 h with PMA, ionomycin; TF1 (early myeloid cell line); and U937 (myelomonocytic cell
30 line) activated with PMA and ionomycin. No signal was detected in non-activated Jurkat, CHA, MRD5, and JY cell lines.

Evaluation of DC subsets: CD34+ progenitors were cultured 6 d with GM-CSF and TNF α , and FACS-sorted into
35 CD1a+ and CD14+ populations. Sorted subsets were cultured 6 more days in GM-CSF and TNF α , and activated with PMA and ionomycin for 1h or 6h. Expression was detected in CD14 derived DC, but not in CD1a derived DC, and the expression was downregulated by PI activation.
40 Much lesser signal was detected in monocytes activated

with PMA and ionomycin; and very weak signals were detected in PBL, both non-activated and PMA, ionomycin activated. No signal was detected in various cells activated with PMA, ionomycin: T cells, granulocytes, or B cells.

Analysis of DC: CD34+ progenitors cultured 12d in GM-CSF and TNF α , non-activated or activated with PMA, ionomycin for either 1 or 6 h. Signals were detected in the CD14 activated subsets, but not in CD1a activated subsets. However, expression was downregulated by PMA, ionomycin activation.

Macrophages were evaluated for expression, and signals were detected in DC (downregulated by PMA, ionomycin activation); monocytes activated with PMA, ionomycin; and PBL (non-activated or activated with PMA, ionomycin).

SDCMP3 expression was not detected by RT-PCR in the following cell types: Langerhans cells, peripheral blood and tonsil CD11c+ or CD11c-negative DC (with or without activation PMA and ionomycin, or IL3 and anti-CD40), B cells (with or without activation PMA and ionomycin, or anti-CD40 mAB), T cells (with or without activation PMA and ionomycin, or anti-CD3 and anti-CD28 mABs).

By sequence expression in cDNA sequence databases, the sequence has been detected in libraries from DC; activated monocytes; and testis tumor.

The murine homolog (1469D4) of SDCMP3 includes a mannose recognition motif (EPN) in its CRD. In addition, the mouse lectin has the consensus WND sequence characteristic of sugar-binding proteins. Accordingly, it can be expected that 1469D4 will have the capacity to bind mannose. As cell walls of microorganisms are rich in mannose, it is possible that antigen-presenting cells (DC) can use the lectin to trap and subsequently degrade microbial antigens through extracellular enzymatic activity.

By analogy to other C-type lectins which exist in closely related forms, it can be predicted that a mannose-binding form of SDCMP3 will be identified from human cells. Such mannose-binding activity on dendritic

cells would represent a target to upregulate for potential benefit in infectious disease treatment. Another possible function of SDCMP3 could be to serve as adhesion molecule between DC and other cell types expressing a ligand, e.g., T cells, thus modulating the immune response.

Sequence homology and chromosomal localization of SDCMP3 strongly suggest that it is a member of a novel C-type lectin family of IRS genes. The sequence of SDCMP3 will be useful to identify other members of the family, by bioinformatics and PCR technology. By analogy to other IRS molecules, SDCMP3 is predicted to associate at the cell surface in a signaling receptor complex. On the basis of its restricted expression in DC and monocytic cells, SDCMP3 would represent a selective target for therapeutic intervention to modulate DC activation. Depending on demonstrated association with an inhibition (ITIM) or activation (ITAM) IRS-signaling pathway, mobilization of SDCMP3 could either suppress or boost immune responses.

In addition, the restricted expression of SDCMP3 suggests the possibility of selective drug delivery to dendritic cells and cells of the monocyte/macrophage series.

Distribution of the mouse SDCMP3 was evaluated by Southern blots from cDNA libraries from various sources. DNA (5 µg) from a primary amplified cDNA library was digested with appropriate restriction enzymes to release the inserts, run on a 1% agarose gel and transferred to a nylon membrane (Schleicher and Schuell, Keene, NH).

Samples for mouse mRNA isolation include: resting mouse fibroblastic L cell line (C200); Braf:ER (Braf fusion to estrogen receptor) transfected cells, control (C201); T cells, TH1 polarized (Mel14 bright, CD4+ cells from spleen, polarized for 7 days with IFN-γ and anti IL-4; T200); T cells, TH2 polarized (Mel14 bright, CD4+ cells from spleen, polarized for 7 days with IL-4 and anti-IFN-γ; T201); T cells, highly TH1 polarized (see Openshaw, et al. (1995) J. Exp. Med. 182:1357-1367; activated with anti-CD3 for 2, 6, 16 h pooled; T202); T

cells, highly TH2 polarized (see Openshaw, et al. (1995) J. Exp. Med. 182:1357-1367; activated with anti-CD3 for 2, 6, 16 h pooled; T203); CD44- CD25+ pre T cells, sorted from thymus (T204); TH1 T cell clone D1.1, resting for 3 weeks after last stimulation with antigen (T205); TH1 T cell clone D1.1, 10 µg/ml ConA stimulated 15 h (T206); TH2 T cell clone CDC35, resting for 3 weeks after last stimulation with antigen (T207); TH2 T cell clone CDC35, 10 µg/ml ConA stimulated 15 h (T208); Mel14+ naive T cells from spleen, resting (T209); Mel14+ T cells, polarized to Th1 with IFN-γ/IL-12/anti-IL-4 for 6, 12, 24 h pooled (T210); Mel14+ T cells, polarized to Th2 with IL-4/anti-IFN-γ for 6, 13, 24 h pooled (T211); unstimulated mature B cell leukemia cell line A20 (B200); unstimulated B cell line CH12 (B201); unstimulated large B cells from spleen (B202); B cells from total spleen, LPS activated (B203); metrizamide enriched dendritic cells from spleen, resting (D200); dendritic cells from bone marrow, resting (D201); monocyte cell line RAW 264.7 activated with LPS 4 h (M200); bone-marrow macrophages derived with GM and M-CSF (M201); macrophage cell line J774, resting (M202); macrophage cell line J774 + LPS + anti-IL-10 at 0.5, 1, 3, 6, 12 h pooled (M203); macrophage cell line J774 + LPS + IL-10 at 0.5, 1, 3, 5, 12 h pooled (M204); aerosol challenged mouse lung tissue, Th2 primers, aerosol OVA challenge 7, 14, 23 h pooled (see Garlisi, et al. (1995) Clinical Immunology and Immunopathology 75:75-83; X206); Nippostrongylus-infected lung tissue (see Coffman, et al. (1989) Science 245:308-310; X200); total adult lung, normal (O200); total lung, rag-1 (see Schwarz, et al. (1993) Immunodeficiency 4:249-252; O205); IL-10 K.O. spleen (see Kuhn, et al. (1991) Cell 75:263-274; X201); total adult spleen, normal (O201); total spleen, rag-1 (O207); IL-10 K.O. Peyer's patches (O202); total Peyer's patches, normal (O210); IL-10 K.O. mesenteric lymph nodes (X203); total mesenteric lymph nodes, normal (O211); IL-10 K.O. colon (X203); total colon, normal (O212); NOD mouse pancreas (see Makino, et al. (1980) Jikken Dobutsu 29:1-13; X205); total thymus, rag-1 (O208); total kidney, rag-1 (O209);

total heart, rag-1 (O202); total brain, rag-1 (O203); total testes, rag-1 (O204); total liver, rag-1 (O206); rat normal joint tissue (O300); and rat arthritic joint tissue (X300).

5 Strong positive signals were detected in: dendritic cells from bone marrow, resting (D201); and bone-marrow macrophages derived with GM and M-CSF (M201). Low signals were detected in total thymus, rag-1 (O208); and total spleen, rag-1 (O207). Barely detectable signals
10 were detected in IL-10 K.O. mesenteric lymph nodes (X203), total adult lung, normal (O200); and total lung, rag-1 (see Schwarz, et al. (1993) Immunodeficiency 4:249-252; O205). Others gave no detectable signal. The high
15 signals suggest that the marker may be useful in distinguishing or characterizing dendritic cell and/or macrophage populations or subpopulations.

The SDCMP4 distribution by PCR: positive signals in: GM-CSF and TNF α treated Dendritic Cells; monocytes
20 activated with PMA and ionomycin; granulocytes activated with PMA and Ionomycin; and PBL (probably Langerhans cells); no detectable signals found in: TF1, Jurkat, MRC5, JY, U937, CHA cell lines; activated T cells; or activated B cells. SDCMP4 is detected in DC (from CD34+ progenitors cultured 12 d in GM-CSF and TNF α), either
25 non-activated or activated with PMA and ionomycin. Signals are also detected in monocytes, granulocytes, and PBL (both non-activated or activated with PMA and ionomycin).

Sequence databases show SDCMP4 sequences in primary
30 dendritic cells (frequent); bone marrow (one); eosinophils (one); placenta subtracted (one); and in T cell lymphoma (two).

The SDCMP3 and SDCMP4 genes display considerable homology with the murine counterpart of human monocyte
35 ASGPR (M-ASGPR). Homology is significant in the carbohydrate-recognition domain which confers specificity to murine monocyte ASGPR for galactose and N-acetylgalactosamine (GalNAc). Sato, et al. (1992) J. Biochem. 111:331-336. In addition, murine monocyte ASGPR
40 has a YENL internalization signal in its cytosolic

domain. A dendrogram of CRD sequences suggests closer relationship of the mouse and human SDCMP3 with the SDCMP2 than with the SDCMP4. These CRDs seem to be more closely related to one another than to the CRD of the hepatic ASGPR.

Murine M-ASGPR functions as a receptor for endocytosis of galactosylated glycoproteins (Ozaki, et al. (1992) J. Biol. Chem. 267:9229-9235), and allows recognition of malignant cells by tumoricidal macrophages (Kawakami, et al. (1994) Jpn. J. Cancer Res. 85:744-749). In this context, murine M-ASGPR was found to be expressed within lung metastatic nodules of mice bearing OV2944-HM-1 metastatic ovarian tumor cells (Imai, et al. (1995) Immunol. 86:591-598). Of interest, human M-ASGPR demonstrates a remarkable specificity for Tn antigen (Suzuki, et al. (1996) J. Immunol. 156:128-135), which bears a cluster of serine or threonine-linked terminal GalNAc, and is associated with human carcinomas (Springer (1989) Mol. Immunol. 26:1-5; and Ørntoft, et al. (1990) Int. J. Cancer 45:666-672).

On the basis of sequence homology, it can be predicted that SDCMPs also function as an endocytic receptor for galactosylated glycoproteins. In addition, ligand internalization via the mannose-receptor, another C-type transmembrane endocytic lectin, results in highly efficient antigen-presentation by DC through the MHC class II pathway. Cella, et al. (1997) Current Opinion Immunol. 9:10-16. By analogy, it is possible that the SDCMPs play a similar role in routing internalized ligands into an antigen-presentation pathway.

Thus, SDCMP4 could be a potential high-efficiency target for loading antigens into DC for enhancing presentation to T cells in immune-based adjuvant therapy. This could be approached by pulsing DC in vitro either with a galactosylated form of antigen, or with anti-SDCMP4 mABs coupled to antigen. In vitro efficiency of presentation could be assayed by activation of antigen-specific T cells. This would focus on presentation of tumor-associated antigens (TAA), due to the inherent therapeutic perspectives of such an approach. Of

particular interest are TAA associated with malignant melanoma.

In addition, the specificity of human M-ASGPR for Tn antigen makes this carcinoma TAA a candidate of choice for targeting the SDCMP4.

As has been recently shown, exogenous antigen can be processed and presented in the MHC class I pathway. See Porgador and Gilboa (1995) J. Exp. Med. 182:255-260; and Paglia, et al. (1996) J. Exp. Med. 183:317-322.

Specialized receptors are likely to perform such a function in DC.

These receptors in DC may be targeted to help produce TAA-specific cytotoxic T cells (CTL), with significant therapeutic potential, as CTL appear to be implicated in the induction of tumor rejection.

XV. Isolation of a binding counterpart

A DC protein can be used as a specific binding reagent, by taking advantage of its specificity of binding, much like an antibody would be used. A binding reagent is either labeled as described above, e.g., fluorescence or otherwise, or immobilized to a substrate for panning methods.

The DC protein is used to screen for a cell line which exhibits binding. Standard staining techniques are used to detect or sort intracellular or surface expressed ligand, or surface expressing transformed cells are screened by panning. Screening of intracellular expression is performed by various staining or immunofluorescence procedures. See also McMahan, et al. (1991) EMBO J. 10:2821-2832.

For example, on day 0, precoat 2-chamber permanox slides with 1 ml per chamber of fibronectin, 10 ng/ml in PBS, for 30 min at room temperature. Rinse once with PBS. Then plate COS cells at $2-3 \times 10^5$ cells per chamber in 1.5 ml of growth media. Incubate overnight at 37° C.

On day 1 for each sample, prepare 0.5 ml of a solution of 66 mg/ml DEAE-dextran, 66 mM chloroquine, and 4 mg DNA in serum free DME. For each set, a positive control is prepared, e.g., of human receptor-FLAG cDNA at

1 and 1/200 dilution, and a negative mock. Rinse cells with serum free DME. Add the DNA solution and incubate 5 hr at 37° C. Remove the medium and add 0.5 ml 10% DMSO in DME for 2.5 min. Remove and wash once with DME. Add 5 1.5 ml growth medium and incubate overnight.

On day 2, change the medium. On days 3 or 4, the cells are fixed and stained. Rinse the cells twice with Hank's Buffered Saline Solution (HBSS) and fix in 4% paraformaldehyde (PFA)/glucose for 5 min. Wash 3X with 10 HBSS. The slides may be stored at -80° C after all liquid is removed. For each chamber, 0.5 ml incubations are performed as follows. Add HBSS/saponin (0.1%) with 32 ml/ml of 1M NaN₃ for 20 min. Cells are then washed with HBSS/saponin 1X. Add protein or protein/antibody 15 complex to cells and incubate for 30 min. Wash cells twice with HBSS/saponin. If appropriate, add first antibody for 30 min. Add second antibody, e.g., Vector anti-mouse antibody, at 1/200 dilution, and incubate for 30 min. Prepare ELISA solution, e.g., Vector Elite ABC 20 horseradish peroxidase solution, and preincubate for 30 min. Use, e.g., 1 drop of solution A (avidin) and 1 drop solution B (biotin) per 2.5 ml HBSS/saponin. Wash cells twice with HBSS/saponin. Add ABC HRP solution and incubate for 30 min. Wash cells twice with HBSS, second 25 wash for 2 min, which closes cells. Then add Vector diaminobenzoic acid (DAB) for 5 to 10 min. Use 2 drops of buffer plus 4 drops DAB plus 2 drops of H₂O₂ per 5 ml of glass distilled water. Carefully remove chamber and rinse slide in water. Air dry for a few minutes, then 30 add 1 drop of Crystal Mount and a cover slip. Bake for 5 min at 85-90° C.

Alternatively, other monocyte protein specific binding reagents are used to affinity purify or sort out cells expressing a receptor. See, e.g., Sambrook, et al. 35 or Ausubel, et al.

Another strategy is to screen for a membrane bound receptor by panning. The receptor cDNA is constructed as described above. The ligand can be immobilized and used to immobilize expressing cells. Immobilization may be 40 achieved by use of appropriate antibodies which

recognize, e.g., a FLAG sequence of a monocyte protein fusion construct, or by use of antibodies raised against the first antibodies. Recursive cycles of selection and amplification lead to enrichment of appropriate clones and eventual isolation of ligand expressing clones.

5

Phage expression libraries can be screened by monocyte protein. Appropriate label techniques, e.g., anti-FLAG antibodies, will allow specific labeling of appropriate clones.

10

Many modifications and variations of this invention can be made without departing from its spirit and scope, as will be apparent to those skilled in the art. The specific embodiments described herein are offered by way of example only, and the invention is to be limited only by the terms of the appended claims, along with the full scope of equivalents to which such claims are entitled.

15

WHAT IS CLAIMED IS:

1. A binding compound comprising an antibody binding site which specifically binds to:
 - a) a primate or rodent SDCMP3 protein; or
 - b) a primate SDCMP4 protein.
2. The binding compound of Claim 1, wherein:
 - a) said antibody binding site is:
 - 1) specifically immunoreactive with a protein of SEQ ID NO: 2 or 4;
 - 2) specifically immunoreactive with a protein of SEQ ID NO: 6 or 8;
 - 3) raised against a purified or recombinantly produced human SDCMP3 protein of SEQ ID NO: 2 or rodent SDCMP3 protein of SEQ ID NO: 4;
 - 4) raised against a purified or recombinantly produced human SDCMP4 protein of SEQ ID NO: 6 or 8;
 - 5) in a monoclonal antibody, Fab, or F(ab)2; or
 - b) said binding compound is:
 - 1) detectably labeled;
 - 2) sterile; or
 - 3) in a buffered composition.
3. A method using the binding compound of Claim 2, comprising contacting said binding compound with a biological sample comprising an antigen to form a binding compound:antigen complex.
4. The method of Claim 3, wherein said biological sample is human, and wherein said binding compound is an antibody.
5. A detection kit comprising said binding compound of Claim 1, and:
 - a) instructional material for the use of said binding compound for said detection; or

b) a compartment providing segregation of said binding compound.

5 6. A substantially pure or isolated polypeptide, which specifically binds to a binding compound of Claim 1.

7. The polypeptide of Claim 6, which:

10 a) comprises at least a fragment of at least 14 amino acid residues from a primate or rodent SDCMP3 protein;

b) comprises at least a fragment of at least 14 amino acid residues from a primate SDCMP4;

c) is a soluble polypeptide;

d) is detectably labeled;

15 e) is in a sterile composition;

f) is in a buffered composition;

g) binds to a sialic acid residue;

h) is recombinantly produced; or

i) has a naturally occurring polypeptide sequence.

20

8. An isolated nucleic acid encoding a polypeptide of Claim 7.

9. The nucleic acid of Claim 8, which:

25 a) comprises at least 30 nucleotides of the coding portion of SEQ ID NO: 1 or 3; or

b) comprises at least 30 nucleotides of the coding portion of SEQ ID NO: 5 or 7.

30

10. A vector comprising the nucleic acid of Claim 8.

11. A cell transfected with said nucleic acid of Claim 8.

35

12. The cell of Claim 11, wherein said nucleic acid consists of the protein encoding portions of SEQ ID NO: 1, 3, or 5.

40

13. A method using at least one strand of said nucleic acid of Claim 8 to form a duplex nucleic acid, said

method comprising a step of contacting said strand to a sample comprising a complementary strand capable of specifically hybridizing.

- 5 14. The method of Claim 13, which:
- a) allows detection of said duplex; or
 - b) allows histological localization of said duplex.
- 10 15. A method of using said binding composition of Claim 1, comprising a step of contacting said binding composition with a sample to form a binding composition:antigen complex.
- 15 16. The method of Claim 15, wherein:
- a) said sample is a biological sample, including a body fluid;
 - b) said antigen is on a cell; or
 - c) said antigen is further purified.
- 20 17. A method using said polypeptide of Claim 6, comprising contacting said polypeptide with a sample to form a binding composition:polypeptide complex.
- 25 18. The method of Claim 17, wherein said polypeptide is further purified.
- 30 19. A method of modulating dendritic cell physiology or function comprising a step of contacting said cell with:
- a) a binding composition of Claim 1;
 - b) a SDCMP3 or SDCMP4 protein; or
 - c) a nucleic acid encoding a SDCMP3 or SDCMP4.
- 35 20. The method of Claim 18, wherein said contacting is in combination with an antigen, including a cell surface, MHC Class I, or MHC Class II antigen.

SEQUENCE LISTING PART

(1) GENERAL INFORMATION:

- (i) APPLICANT: SCHERING CORPORATION
- (ii) TITLE OF INVENTION: Isolated Mammalian Membrane Protein Genes and Related Reagents
- (iii) NUMBER OF SEQUENCES: 8
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Schering Corporation
 - (B) STREET: 2000 Galloping Hill Road
 - (C) CITY: Kenilworth
 - (D) STATE: New Jersey
 - (E) COUNTRY: USA
 - (F) ZIP: 07033-0530
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: Apple Macintosh
 - (C) OPERATING SYSTEM: Macintosh 7.1
 - (D) SOFTWARE: Microsoft Word 6.0
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME:
 - (B) REGISTRATION NUMBER:
 - (C) REFERENCE/DOCKET NUMBER:
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (908)298 4000
 - (B) TELEFAX: (908)298 5388

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 850 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 108..593
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

2

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 162 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Met Gln Glu Gln Gln Pro Gln Ser Thr Glu Lys Arg Gly Trp Leu
 1             5             10             15
Ser Leu Arg Leu Trp Ser Val Ala Gly Ile Ser Ile Ala Leu Leu Ser
      20             25             30
Ala Cys Phe Ile Val Ser Cys Val Val Thr Tyr His Phe Thr Tyr Gly
      35             40             45
Glu Thr Gly Lys Arg Leu Ser Glu Leu His Ser Tyr His Ser Ser Leu
      50             55             60
Thr Cys Phe Ser Glu Gly Thr Lys Val Pro Ala Trp Gly Cys Cys Pro
      65             70             75             80
Ala Ser Trp Lys Ser Phe Gly Ser Ser Cys Tyr Phe Ile Ser Ser Glu
      85             90             95
Glu Lys Val Trp Ser Lys Ser Glu Gln Asn Cys Val Glu Met Gly Ala
      100            105            110
His Leu Val Val Phe Asn Thr Glu Ala Glu Gln Asn Phe Ile Val Gln
      115            120            125
Gln Leu Asn Glu Ser Phe Ser Tyr Phe Leu Gly Leu Ser Asp Pro Gln
      130            135            140
Gly Asn Asn Asn Trp Gln Trp Ile Asp Lys Thr Pro Tyr Glu Lys Asn
      145            150            155            160
Val Arg

```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 630 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..627

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATG GTG CAG GAA AGA CAA TCC CAA GGG AAG GGA GTC TGC TGG ACC CTG	48
Met Val Gln Glu Arg Gln Ser Gln Gly Lys Gly Val Cys Trp Thr Leu	
1 5 10 15	
AGA CTC TGG TCA GCT GCT GTG ATT TCC ATG TTA CTC TTG AGT ACC TGT	96
Arg Leu Trp Ser Ala Ala Val Ile Ser Met Leu Leu Leu Ser Thr Cys	
20 25 30	
TTC ATT GCG AGC TGT GTG GTG ACT TAC CAA TTT ATT ATG GAC CAG CCC	144
Phe Ile Ala Ser Cys Val Val Thr Tyr Gln Phe Ile Met Asp Gln Pro	
35 40 45	
AGT AGA AGA CTA TAT GAA CTT CAC ACA TAC CAT TCC AGT CTC ACC TGC	192
Ser Arg Arg Leu Tyr Glu Leu His Thr Tyr His Ser Ser Leu Thr Cys	
50 55 60	
TTC ACT GAA GGG ACT ATG GTG TCA GAA AAA ATG TGG GGA TGC TGC CCA	240
Phe Ser Glu Gly Thr Met Val Ser Glu Lys Met Trp Gly Cys Cys Pro	
65 70 75 80	
AAT CAC TGG AAG TCA TTT GGC TCC AGC TGC TAC CTC ATT TCT ACC AAG	288
Asn His Trp Lys Ser Phe Gly Ser Ser Cys Tyr Leu Ile Ser Thr Lys	
85 90 95	
GAG AAC TTC TGG AGC ACC AGT GAG CAG AAC TGT GTT CAG ATG GGG GCT	336
Glu Asn Phe Trp Ser Thr Ser Glu Gln Asn Cys Val Gln Met Gly Ala	
100 105 110	
CAT CTG GTG GTG ATC AAT ACT GAA GCG GAG CAG AAT TTC ATC ACC CAG	384
His Leu Val Val Ile Asn Thr Glu Ala Glu Gln Asn Phe Ile Thr Gln	
115 120 125	
CAG CTG AAT GAG TCA CTT TCT TAC TTC CTG GGT CTT TCG GAT CCA CAA	432
Gln Leu Asn Glu Ser Leu Ser Tyr Phe Leu Gly Leu Ser Asp Pro Gln	
130 135 140	
GGT AAT GGC AAA TGG CAA TGG ATC GAT GAT ACT CCT TTC AGT CAA AAT	480
Gly Asn Gly Lys Trp Gln Trp Ile Asp Asp Thr Pro Phe Ser Gln Asn	
145 150 155 160	
GTC AGG TTC TGG CAC CCC CAT GAA CCC AAT CTT CCA GAA GAG CGG TGT	528
Val Arg Phe Trp His Pro His Glu Pro Asn Leu Pro Glu Glu Arg Cys	
165 170 175	
GTT TCA ATA GTT TAC TGG AAT CCT TCG AAA TGG GGC TGG AAT GAT GTT	576
Val Ser Ile Val Tyr Trp Asn Pro Ser Lys Trp Gly Trp Asn Asp Val	
180 185 190	
TTC TGT GAT AGT AAA CAC AAT TCA ATA TGT GAA ATG AAG AAG ATT TAC	624
Phe Cys Asp Ser Lys His Asn Ser Ile Cys Glu Met Lys Lys Ile Tyr	
195 200 205	
CTA TGA	630
Leu	

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 209 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

Met Val Gln Glu Arg Gln Ser Gln Gly Lys Gly Val Cys Trp Thr Leu
 1             5             10             15

Arg Leu Trp Ser Ala Ala Val Ile Ser Met Leu Leu Leu Ser Thr Cys
      20             25             30

Phe Ile Ala Ser Cys Val Val Thr Tyr Gln Phe Ile Met Asp Gln Pro
      35             40             45

Ser Arg Arg Leu Tyr Glu Leu His Thr Tyr His Ser Ser Leu Thr Cys
      50             55             60

Phe Ser Glu Gly Thr Met Val Ser Glu Lys Met Trp Gly Cys Cys Pro
      65             70             75             80

Asn His Trp Lys Ser Phe Gly Ser Ser Cys Tyr Leu Ile Ser Thr Lys
      85             90             95

Glu Asn Phe Trp Ser Thr Ser Glu Gln Asn Cys Val Gln Met Gly Ala
      100            105            110

His Leu Val Val Ile Asn Thr Glu Ala Glu Gln Asn Phe Ile Thr Gln
      115            120            125

Gln Leu Asn Glu Ser Leu Ser Tyr Phe Leu Gly Leu Ser Asp Pro Gln
      130            135            140

Gly Asn Gly Lys Trp Gln Trp Ile Asp Asp Thr Pro Phe Ser Gln Asn
      145            150            155            160

Val Arg Phe Trp His Pro His Glu Pro Asn Leu Pro Glu Glu Arg Cys
      165            170            175

Val Ser Ile Val Tyr Trp Asn Pro Ser Lys Trp Gly Trp Asn Asp Val
      180            185            190

Phe Cys Asp Ser Lys His Asn Ser Ile Cys Glu Met Lys Lys Ile Tyr
      195            200            205

Leu

```

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1018 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 160..900

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATCTGGTTGA ACTACTTAAG CTTAATTTGT TAACTCCGG TAAGTACCTA GCCCACATGA	60
TTTGACTCAG AGATTCTCTT TTGTCCACAG ACAGTCATCT CAGGAGCAGA AAGAAAAGAG	120
CTCCCAAATG CTATATCTAT TCAGGGGCTC TCAAGAACA ATG GAA TAT CAT CCT	174
Met Glu Tyr His Pro	
1 5	
GAT TTA GAA AAT TTG GAT GAA GAT GGA TAT ACT CAA TTA CAC TTC GAC	222
Asp Leu Glu Asn Leu Asp Glu Asp Gly Tyr Thr Gln Leu His Phe Asp	
10 15 20	
TCT CAA AGC AAT ACC AGG ATA GCT GTT GTT TCA GAG AAA GGA TCG TGT	270
Ser Gln Ser Asn Thr Arg Ile Ala Val Val Ser Glu Lys Gly Ser Cys	
25 30 35	
GCT GCA TCT CCT CCT TGG CGC CTC ATT GCT GTA ATT TTG GGA ATC CTA	318
Ala Ala Ser Pro Pro Trp Arg Leu Ile Ala Val Ile Leu Gly Ile Leu	
40 45 50	
TGC TTG GTA ATA CTG GTG ATA GCT GTG GTC CTG GGT ACC ATG GCT ATT	366
Cys Leu Val Ile Leu Val Ile Ala Val Val Leu Gly Thr Met Ala Ile	
55 60 65	
TGG AGA TCC AAT TCA GGA AGC AAC ACA TTG GAG AAT GGC TAC TTT CTA	414
Trp Arg Ser Asn Ser Gly Ser Asn Thr Leu Glu Asn Gly Tyr Phe Leu	
70 75 80 85	
TCA AGA AAT AAA GAG AAC CAC AGT CAA CCC ACA CAA TCA TCT TTA GAA	462
Ser Arg Asn Lys Glu Asn His Ser Gln Pro Thr Gln Ser Ser Leu Glu	
90 95 100	
GAC AGT GTG ACT CCT ACC AAA GCT GTC AAA ACC ACA GGG GTT CTT TCC	510
Asp Ser Val Thr Pro Thr Lys Ala Val Lys Thr Thr Gly Val Leu Ser	
105 110 115	
AGC CCT TGT CCT CCT AAT TGG ATT ATA TAT GAG AAG AGC TGT TAT CTA	558
Ser Pro Cys Pro Pro Asn Trp Ile Ile Tyr Glu Lys Ser Cys Tyr Leu	
120 125 130	

TTC AGC ATG TCA CTA AAT TCC TGG GAT GGA AGT AAA AGA CAA TGC TGG	606
Phe Ser Met Ser Leu Asn Ser Trp Asp Gly Ser Lys Arg Gln Cys Trp	
135 140 145	
CAA CTG GGC TCT AAT CTC CTA AAG ATA GAC AGC TCA AAT GAA TTG GGA	654
Gln Leu Gly Ser Asn Leu Leu Lys Ile Asp Ser Ser Asn Glu Leu Gly	
150 155 160 165	
TTT ATA GTA AAA CAA GTG TCT TCC CAA CCT GAT AAT TCA TTT TGG ATA	702
Phe Ile Val Lys Gln Val Ser Ser Gln Pro Asp Asn Ser Phe Trp Ile	
170 175 180	
GGC CTT TCT CGG CCC CAG ACT GAG GTA CCA TGG CTC TGG GAG GAT GGA	750
Gly Leu Ser Arg Pro Gln Thr Glu Val Pro Trp Leu Trp Glu Asp Gly	
185 190 195	
TCA ACA TTC TCT TCT AAC TTA TTT CAG ATC AGA ACC ACA GCT ACC CAA	798
Ser Thr Phe Ser Ser Asn Leu Phe Gln Ile Arg Thr Thr Ala Thr Gln	
200 205 210	
GAA AAC CCA TCT CCA AAT TGT GTA TGG ATT CAC GTG TCA GTC ATT TAT	846
Glu Asn Pro Ser Pro Asn Cys Val Trp Ile His Val Ser Val Ile Tyr	
215 220 225	
GAC CAA CTG TGT AGT GTG CCC TCA TAT AGT ATT TGT GAG AAG AAG TTT	894
Asp Gln Leu Cys Ser Val Pro Ser Tyr Ser Ile Cys Glu Lys Lys Phe	
230 235 240 245	
TCA ATG TAAGGGGAAG GGTGGAGAAG GAGAGAGAAA TATGTGAGGT AGTTAAGGAG	950
Ser Met	
GACAGAAAAC AGAACAGAAA AGAGTAACAG CTGAGGGTCA AGATAAATGC AGAAAATGTT	1010
TAGAGAGC	1018

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 247 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Glu Tyr His Pro Asp Leu Glu Asn Leu Asp Glu Asp Gly Tyr Thr	
1 5 10 15	
Gln Leu His Phe Asp Ser Gln Ser Asn Thr Arg Ile Ala Val Val Ser	
20 25 30	
Glu Lys Gly Ser Cys Ala Ala Ser Pro Pro Trp Arg Leu Ile Ala Val	
35 40 45	

```

Ile Leu Gly Ile Leu Cys Leu Val Ile Leu Val Ile Ala Val Val Leu
 50                      55                      60

Gly Thr Met Ala Ile Trp Arg Ser Asn Ser Gly Ser Asn Thr Leu Glu
 65                      70                      75                      80

Asn Gly Tyr Phe Leu Ser Arg Asn Lys Glu Asn His Ser Gln Pro Thr
                      85                      90                      95

Gln Ser Ser Leu Glu Asp Ser Val Thr Pro Thr Lys Ala Val Lys Thr
                      100                      105                      110

Thr Gly Val Leu Ser Ser Pro Cys Pro Pro Asn Trp Ile Ile Tyr Glu
                      115                      120                      125

Lys Ser Cys Tyr Leu Phe Ser Met Ser Leu Asn Ser Trp Asp Gly Ser
                      130                      135                      140

Lys Arg Gln Cys Trp Gln Leu Gly Ser Asn Leu Leu Lys Ile Asp Ser
                      145                      150                      155                      160

Ser Asn Glu Leu Gly Phe Ile Val Lys Gln Val Ser Ser Gln Pro Asp
                      165                      170                      175

Asn Ser Phe Trp Ile Gly Leu Ser Arg Pro Gln Thr Glu Val Pro Trp
                      180                      185                      190

Leu Trp Glu Asp Gly Ser Thr Phe Ser Ser Asn Leu Phe Gln Ile Arg
                      195                      200                      205

Thr Thr Ala Thr Gln Glu Asn Pro Ser Pro Asn Cys Val Trp Ile His
                      210                      215                      220

Val Ser Val Ile Tyr Asp Gln Leu Cys Ser Val Pro Ser Tyr Ser Ile
                      225                      230                      235                      240

Cys Glu Lys Lys Phe Ser Met
                      245

```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 880 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 160..762

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

ATCTGGTTGA ACTACTTAAG CTTAATTTGT TAAACTCCGG TAAGTACCTA GCCCACATGA	60
TTTGACTCAG AGATTCTCTT TTGTCCACAG ACAGTCATCT CAGGAGCCGA AAGAAAAGAG	120
CTCCCAAATG CTATATCTAT TCAGGGGCTC TCAAGAACA ATG GAA TAT CAT CCT	174
Met Glu Tyr His Pro	
1 5	
GAT TTA GAA AAT TTG GAT GAA GAT GGA TAT ACT CAA TTA CAC TTC GAC	222
Asp Leu Glu Asn Leu Asp Glu Asp Gly Tyr Thr Gln Leu His Phe Asp	
10 15 20	
TCT CAA AGC AAT ACC ATG ATA GCT GTT GTT TCA GAG AAA GGA TCG TGT	270
Ser Gln Ser Asn Thr Met Ile Ala Val Val Ser Glu Lys Gly Ser Cys	
25 30 35	
GCT GCA TCT CCT CCT TGG CGC CTC ATT GCT GTA ATT TTG GGA ATC CTA	318
Ala Ala Ser Pro Pro Trp Arg Leu Ile Ala Val Ile Leu Gly Ile Leu	
40 45 50	
TGC TTG GTA ATA CTG GTG ATA GCT GTG GTC CTG GGT ACC ATG GGG GTT	366
Cys Leu Val Ile Leu Val Ile Ala Val Val Leu Gly Thr Met Gly Val	
55 60 65	
CTT TCC AGC CCT TGT CCT CCT AAT TGG ATT ATA TAT GAG AAG AGC TGT	414
Leu Ser Ser Pro Cys Pro Pro Asn Trp Ile Ile Tyr Glu Lys Ser Cys	
70 75 80 85	
TAT CTA TTC AGC ATG TCA CTA AAT TCC TGG GAT GGA AGT AAA AGA CAA	462
Tyr Leu Phe Ser Met Ser Leu Asn Ser Trp Asp Gly Ser Lys Arg Gln	
90 95 100	
TGC TGG CAA CTG GGC TCT AAT CTC CTA AAG ATA GAC AGC TCA AAT GAA	510
Cys Trp Gln Leu Gly Ser Asn Leu Leu Lys Ile Asp Ser Ser Asn Glu	
105 110 115	
TTG GGA TTT ATA GTA AAA CAA GTG TCT TCC CAA CCT GAT AAT TCA TTT	558
Leu Gly Phe Ile Val Lys Gln Val Ser Ser Gln Pro Asp Asn Ser Phe	
120 125 130	
TGG ATA GGC CTT TCT CGG CCC CAG ACT GAG GTA CCA TGG CTC TGG GAG	606
Trp Ile Gly Leu Ser Arg Pro Gln Thr Glu Val Pro Trp Leu Trp Glu	
135 140 145	
GAT GGA TCA ACA TTC TCT TCT AAC TTA TTT CAG ATC AGA ACC ACA GCT	654
Asp Gly Ser Thr Phe Ser Ser Asn Leu Phe Gln Ile Arg Thr Thr Ala	
150 155 160 165	
ACC CAA GAA AAC CCA TCT CCA AAT TGT GTA TGG ATT CAC GTG TCA GTC	702
Thr Gln Glu Asn Pro Ser Pro Asn Cys Val Trp Ile His Val Ser Val	
170 175 180	
ATT TAT GAC CAA CTG TGT AGT GTG CCC TCA TAT AGT ATT TGT GAC AAC	750
Ile Tyr Asp Gln Leu Cys Ser Val Pro Ser Tyr Ser Ile Cys Glu Lys	
185 190 195	

AAG TTT TCA ATG TAAGGGGAAG GGTGGAGAAG GAGAGAGAAA TATGTGAGGT 802
 Lys Phe Ser Met
 200

AGTTAAGGAG GACAGAAAAC AGAACAGAAA AGAGTAACAG CTGAGGGTCA AGATAAATGC 862

AGAAAATGTT TAGAGAGC 880

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 201 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Glu Tyr His Pro Asp Leu Glu Asn Leu Asp Glu Asp Gly Tyr Thr
 1 5 10 15

Gln Leu His Phe Asp Ser Gln Ser Asn Thr Met Ile Ala Val Val Ser
 20 25 30

Glu Lys Gly Ser Cys Ala Ala Ser Pro Pro Trp Arg Leu Ile Ala Val
 35 40 45

Ile Leu Gly Ile Leu Cys Leu Val Ile Leu Val Ile Ala Val Val Leu
 50 55 60

Gly Thr Met Gly Val Leu Ser Ser Pro Cys Pro Pro Asn Trp Ile Ile
 65 70 75 80

Tyr Glu Lys Ser Cys Tyr Leu Phe Ser Met Ser Leu Asn Ser Trp Asp
 85 90 95

Gly Ser Lys Arg Gln Cys Trp Gln Leu Gly Ser Asn Leu Leu Lys Ile
 100 105 110

Asp Ser Ser Asn Glu Leu Gly Phe Ile Val Lys Gln Val Ser Ser Gln
 115 120 125

Pro Asp Asn Ser Phe Trp Ile Gly Leu Ser Arg Pro Gln Thr Glu Val
 130 135 140

Pro Trp Leu Trp Glu Asp Gly Ser Thr Phe Ser Ser Asn Leu Phe Gln
 145 150 155 160

Ile Arg Thr Thr Ala Thr Gln Glu Asn Pro Ser Pro Asn Cys Val Trp
 165 170 175

Ile His Val Ser Val Ile Tyr Asp Gln Leu Cys Ser Val Pro Ser Tyr
 180 185 190

WO 99/47673

PCT/US99/03740

Ser	Ile	Cys	Glu	Lys	Lys	Phe	Ser	Met
		195					200	

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/03740

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/12 C07K14/705 C07K16/28 G01N33/68 C12N1/21
C12N5/10 C12N1/19 C12Q1/68 A61K39/395 A61K38/17
A61K31/70 //C12N9/56

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 02456 A (COCKS BENJAMIN GRAEME ; INCYTE PHARMA INC (US); AU YOUNG JANICE (US) 22 January 1998 (1998-01-22) the whole document ---	1-20
P, X	WO 98 28332 A (TAKASHIMA AKIRA ; UNIV TEXAS (US); ARIIZUMI KIYOSHI (US)) 2 July 1998 (1998-07-02) the whole document --- -/--	1-20

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

29 September 1999

Date of mailing of the international search report

12/10/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040. Tx. 31 651 epo nl.
Fax: (+31-70) 340-3016

Authorized officer

Van der Schaal, C

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/03740

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ARIIZUMI, KIYOSHI ET AL: "Identification of dectin -2, a second member of a new dendritic cell-associated C-type lectin family." JOURNAL OF INVESTIGATIVE DERMATOLOGY, (1997) VOL. 108, NO. 4, PP. 564. MEETING , XP002116889 see abstract 159	1-12, 15-18,20
O,X	& ANNUAL MEETING OF THE SOCIETY FOR INVESTIGATIVE DERMATOLOGY WASHINGTON, D.C., USA APRIL 23-27, 1997 ---	
X	ARIIZUMI, K. ET AL: "Identification and characterization of genomic DNA for dectin -2." JOURNAL OF INVESTIGATIVE DERMATOLOGY, (1997) VOL. 109, NO. 2, PP. 254. MEETING, XP002116890 see abstract 3	1-12, 15-20
O,X	& FIFTH INTERNATIONAL WORKSHOP ON LANGERHANS CELLS SALZBURG, AUSTRIA SEPTEMBER 5-7, 1997 ---	
X	YOKOTA, K. ET AL: "Identification of human equivalents of murine dectin -1 and dectin -2." JOURNAL OF INVESTIGATIVE DERMATOLOGY, (1997) VOL. 109, NO. 2, PP. 267., August 1997 (1997-08), XP002116891 see abstract84	1-18,20
O,X	& FIFTH INTERNATIONAL WORKSHOP ON LANGERHANS CELLS SALZBURG, AUSTRIA SEPTEMBER 5-7, 1997 ---	
X	YOKOTA, K. ET AL: "Molecular cloning of a new human C-type lectin with significant homology to mouse dectin -2." JOURNAL OF DERMATOLOGICAL SCIENCE, (MARCH, 1998) VOL. 16, NO. SUPPL. 1, PP. S89., XP002116892 see abstract 0531	1-4,6-12
P,O, X	& THIRD JOINT MEETING OF THE EUROPEAN SOCIETY FOR DERMATOLOGICAL RESEARCH, JAPANESE SOCIETY FOR INVESTIGATIVE DERMATOLOGY, SOCIETY FOR INVESTIGATIVE DERMATOLOGY, COLOGNE MAY 7-10 1998 ---	
	--- -/--	

INTERNATIONAL SEARCH REPORT

International Application No

PC1/US 99/03740

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>ARIIZUMI, K. (1) ET AL: "Isolation of a dendritic cell-specific transcription unit, dectin -2 promoter." JOURNAL OF DERMATOLOGICAL SCIENCE, (MARCH, 1998) VOL. 16, NO. SUPPL. 1, PP. S3. MEETING INFO.: THIRD JOINT MEETING OF THE EUROPEAN SOCIETY FOR DERMATOLOGICAL RESEARCH, JAPANESE SOCIETY FOR INVESTIGATIVE DERMATOLOGY, COLOGNE MAY 7-10 1998, XP002116893 see abstract 0015</p>	6-12
P,O, X	<p>& THIRD JOINT MEETING OF THE EUROPEAN SOCIETY FOR DERMATOLOGICAL RESEARCH, JAPANESE SOCIETY FOR INVESTIGATIVE DERMATOLOGY, SOCIETY FOR INVESTIGATIVE DERMATOLOGY, COLOGNE MAY 7-10 1998 ---</p>	
P,X	<p>DATABASE WPI Section Ch, Week 9911 Derwent Publications Ltd., London, GB; Class B04, AN 99-125479 XP002116016 & JP 11 001497 A (TAKEDA CHEM IND LTD), 6 January 1999 (1999-01-06) abstract -& EMBL/GENBANK DATABASES Accession no X01259 31 March 1999 "Human DC3 protein coding sequence" XP002116017 -& EMBL/GENBANK DATABASES Accession no W73888 31 March 1999 "Human DC3 protein sequence" XP002116018 -& EMBL/GENBANK DATABASES Accession no X01260 31 March 1999 "Human DC3' protein coding sequence" XP002116019 -& EMBL/GENBANK DATABASES Accession no W73889 31 March 1999 "Human DC3' protein sequence" XP002116020 ---</p>	1-18,20
P,A	<p>WO 99 02562 A (SCHERING CORP) 21 January 1999 (1999-01-21) -----</p>	

INTERNATIONAL SEARCH REPORT

national application No.

PCT/US 99/ 03740

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Remark: Although claim 19 is partially directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

See additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☒ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-20 partially

SDCMP3, its encoding nucleotides and specific antibodies against SDCMP3 and their uses

2. Claims: 1-20 partially

SDCMP4, its encoding nucleotides and specific antibodies against SDCMP4 and their uses

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 99/03740

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9802456 A	22-01-1998	US 5871964 A AU 3665097 A	16-02-1999 09-02-1999
WO 9828332 A	02-07-1998	AU 5805698 A	17-07-1998
JP 11001497 A	06-01-1999	NONE	
WO 9902562 A	21-01-1999	AU 8271298 A	08-02-1999